

THÈSES

PRÉSENTÉES

A LA FACULTÉ DES SCIENCES
DE L'UNIVERSITÉ DE PARIS

POUR OBTENIR

LE GRADE DE DOCTEUR ÈS SCIENCES MATHÉMATIQUES

PAR

HAYRI KOREZLIOGLU

1^{re} THÈSE

PRÉVISION ET FILTRAGE LINÉAIRES, APPLICATIONS
A LA REPRÉSENTATION CANONIQUE DES PROCESSUS ALÉATOIRES DU SECOND ORDRE
ET A LA THÉORIE DE L'INFORMATION

2^e THÈSE

PROPOSITIONS DONNÉES PAR LA FACULTÉ

Soutenues, le 7 mars 1963, devant la Commission d'Examen

Jury	{	MM. J. A. VILLE	Président.
		R. FORTET	Examineurs
		Ch. PISOT	

Prévision et filtrage linéaires, applications
à la représentation canonique des processus aléatoires
du second ordre et à la théorie de l'information





FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE PARIS

Doyen M. ZAMANSKY

PROFESSEURS

G. JULIA T	Analyse supérieure et algèbre sup.	COUTEAUX T	Biologie animale (P.C.B.).
DE BROGLIE T	Théories physiques.	MAY T	Zoologie (Orsay).
PRENANT T	Anatomie et histologie comparées.	CHOQUET T	Théorie des fonctions et topologie.
GRASSÉ T	Evolution des êtres organisés.	FELDMANN T	Biologie végétale marine.
PRÉVOST T	Chimie organique.	GUINIER T	Physique des solides (Orsay).
CHAUDRON T	Chimie appliquée.	JOST T	Physiologie comparée.
WYART T	Minéralogie et cristal.	FORTET T	Calcul des probabilités et physique mathématique.
TEISSIER T	Zoologie.	SCHWARTZ T	Calcul diff. et intégral.
MANGENOT T	Biologie vég. (Orsay).	CHOUARD T	Physiologie végétale.
AUGER T	Physique quantique et relativité.	MALAVARD T	Aviation (Techn. aéronaut.).
MONNIER T	Physiologie générale.	BRELOT T	Calcul diff. et intégral.
PIVETEAU T	Paléontologie.	NORMANT T	Synthèse organique générale.
ROCARD T	Physique (E.N.S.).	BÉNARD T	Chimie minérale.
CARTAN T	Mathématiques (E.N.S.).	BUVAT T	Botanique (E.N.S.).
LAFFITTE T	Chimie générale.	DUGUÉ T	Statistiques mathématiques.
FAVARD T	Géométrie supérieure.	SOULAIRAC T	Psychophysiologie.
COULOMB T	Physique du Globe.	ULRICH T	Physiologie végétale.
Mlle COUSIN T	Biologie animale (S.P.C.N.).	MARÉCHAL T	Optique théorique et appliquée.
CHRÉTIEN T	Chimie minérale.	KIRRMANN T	Théories chimiques.
DRACH T	Zoologie.	CHADEFAUD T	Botanique.
KASTLER T	Physique (E.N.S.).	Mlle LE BRETON. T	Physiologie de la nutrition.
EPHRUSSI T	Génétique.	SALEM T	Mathématiques générales.
RIVIÈRE T	Géologie S.P.C.N. et sédim.	LELONG T	Applic. de l'analyse à la géométrie.
GAUTHERET T	Biologie vég. (P.C.B.).	DEVILLERS T	Anatomie et histologie comp.
LUCAS R. T	Recherches physiques.	EHRESMANN T	Topologie algébrique.
THOMAS A. T	Biologie cellulaire.	FRANÇON T	Physique (S.P.C.N.).
ARNULF T	Optique appliquée.	GLANGEAUD T	Géographie phys. et géologie dynamique.
MORAND T	Physique Enseignement.	GODEMENT T	Mathématiques (M.P.C.).
SOLEILLET T	Physique (P.C.B.).	PISOT T	Tech. mathém. de la phys.
FORTIER T	Mécanique exp. des fluides.	ROCH T	Géologie.
DANJON T	Astronomie.	SCHATZMAN T	Astrophysique.
PETIT T	Biologie maritime.	TERMIER T	Paléontologie stratig.
QUENEY T	Météorologie et dynamique atmosphérique.	ZAMANSKY T	Mathématiques générales.
GALLIEN T	Embryologie.	LENNUIER T	Physique (M.G.P.).
EICHHORN T	Botanique.	RIZET T	Génétique (Orsay).
DE CUGNAC T	Biologie végétale (S.P.C.N.).	ROUTHIER T	Géologie appliquée.
Mlle CAUCHOIS .. T	Chimie physique.	Mme TONNELAT.. T	Physique théorique.
THELLIER T	Physique du Globe.	DIXMIER T	Mathématiques (M.P.C.).
L'HÉRITIER T	Biologie générale (Orsay).	SOUCHAY T	Chimie générale.
GRIVET T	Radioélectricité (Orsay).	AIGRAIN T	Electrotechnique générale.
PONCIN T	Mécanique générale.	BRUSSET T	Chimie systématique.
DUBREIL T	Arithmétique et théorie des nombres.	M. LÉVY T	Physique théorique des hautes énergies (Orsay).
QUELET T	Chimie organique.	Mme CHAIX T	Chimie biologique.
CAGNIARD T	Géophysique appliquée.	Mme HUREL-PY. T	Biologie végétale (P.C.B.).
CHAMPETIER T	Chimie macromoléculaire.	PIAUX T	Chimie (M.P.C.).
CUVILLIER T	Micropaléontologie.	BRUN T	Mécan. théorique des fluides.
JUNG T	Pétrographie	LEDERER T	Chimie biologique.
TRILLAT T	Microscopie et diffraction électronique.	Mme DUBREIL .. T	Mathém. (Agrégation).
WIEMANN T	Chimie organique et struc.	Mme LELONG T	Mathématiques (E.N.S.).
JACQUINOT T	Spectroscopie et physique céleste (Orsay).	BELLAIR T	Géologie (S.P.C.N.).
VASSY T	Physique de l'atmosphère.	COTTE T	Physique électricité.
DESTOUCHES T	Théories physiques.	DUBOIS J.E. T	Chimie organique.
AMIEL T	Chimie générale.	LAMOTTE T	Zoologie (E.N.S.).
HOCART T	Minéralogie et cristallog.	MICHEL A. T	Chimie minérale (Orsay).
J.P. MATHIEU ... T	Physique (Optique).	OLMER T	Energétique générale.
		ROUAULT T	Electricité (Orsay).

GAUTHIER T Mécanique appliquée.
 BARCHEWITZ T Chimie physique.
 BROSSEL T Physique atomique.
 BUSER T Physiologie comparée.
 CAMUS T Physiologie vég. (B.M.P.V.).
 CASTAING T Physique (Orsay).
 CURIEN T Minéralo. et cristallog.
 MOYSE T Physiologie vég. (Orsay).
 NOIROT T Evolution des êtres org.
 PANNETIER T Chimie générale (Orsay).
 POSSOMPES T Zoologie.
 PULLMANN T Chimie théorique.
 TEILLAC T Physique nucl. et radioactivité.
 TONNELAT T Biologie physico-chim.
 VILLE T Econométrie
 WILLEMART T Chimie (P.C.B.).
 DODE T Chimie (Orsay).
 FREYMAN T Recherches physiques.
 GUINOCHET T Biol. vég. (S.P.C.N.) (Orsay).
 ROLLET T Chimie (P.C.B.).
 Mlle JOSIEN T Chimie (S.P.C.N.).
 CHEVALLEY T Géométrie algébrique et théorie des groupes.
 ARNOULT T Elect onique (Orsay).
 CHAPELLE T Physique (Orsay).
 DELANGE T Mathématiques (Orsay).
 DENY T Mathématiques gén. (Orsay).
 GERMAIN T Mécan. théor. des fluides.
 LUCAS G. T Géologie.
 ALLARD T Chimie physique.
 BERTHELOT T Phys. fondamentale des hautes énergies (Orsay).
 BRICART T Météorologie et phys. de l'atmosphère.
 Mme ALBE-
 FESSARD T Psychophysiologie.
 FRIEDEL T Physique des solides (Orsay).
 M. JULIA Etudes des moléculaires naturelles complexes
 LENDER T Biol. animale (S.P.C.N.) (Orsay).
 MAGAT Physico-chimie des radio-éléments.
 Mlle QUINTIN ... T Electrochimie
 MONOD T Chimie du métabolisme.
 BENOIT T Physiol. anim. (Orsay).
 DE POSSEL T Analyse numérique.
 CHARLOT T Chimie analytique.
 HAISSINSKY T Radiochimie.
 LEMÉE T Physiol. vég. B.M.P.V. (Orsay).
 GUÉRIN T Chimie M.P.C. (Orsay).
 CHATELET M. ... T Chimie (P.C.B.).
 JEAN T Physique nucléaire (Orsay).
 MATTLER T Physique S.P.C.N. (O.).

MICHEL L. Phys. théor. des hautes énergies (Orsay).
 BERGERARD T Zoologie (Orsay).
 BERTIN T Electronique (Orsay).
 BOUSSIÈRES Chimie des radioéléments (Orsay).
 Mme COUTURE.. T Thermodynamique et méc. physique.
 ELLENBERGER ... T Géologie appli. s/ le terr.
 FRANC Biol. animale (P.C.B.).
 SCHNELL Botanique tropicale (Orsay).
 STOLKOWSKI Physiologie générale.
 ACHER Chimie biol. (Orsay).
 BAUDOIN Biologie animale (P.C.B.).
 BEAUMONT Biologie animale (S.P.C.N.).
 BLAMONT T Géophysique ionosphère.
 BOURREAU Botanique.
 CURIE D. Physique (M.G.P.).
 DURAND-DELGA ... Géologie.
 FRÉON T Chimie organique (Orsay).
 HELLER T Physiologie végétale.
 TORTRAT T Calcul des probabilités.
 CAILLEUX Géologie (S.P.C.N.).
 MAGNAN T Physique (S.P.C.N.).
 DAUDEL T Mécanique ondulatoire appl. à la chim. théor. et à la phys. nucléaire.
 Mlle FOURCROY.. Biol. végétal: (S.P.C.N.).
 Mme CHOQUET-
 BRUHAT T Méc. analytique et méc. céleste.
 Mlle VEIL T Physiologie générale.
 LWOFF T Microbiologie
 SIESTRUNCK T Mécan. phys et expérim.
 LACOMBE T Métallurgie des métaux nucl.
 BROCHARD T Physique P.C.B. (Orsay).
 CABANNES T Mécanique générale.
 LESIEUR T Mathématiques M.P.C. (Orsay).
 MAZET T Mécanique générale (Orsay).
 CAIRE Géographie phys. et géol. dynamique.
 COMOLET T Mécanique des fluides.
 CORABŒUF Physiologie animale.
 DELLOUE Physique de l'exosphère.
 GORENFLOT Botanique II (Orsay).
 HOUILLO Biologie animale (P.C.B.).
 JULLIEN J. Chimie (P.C.B.).
 MALGRANGE T Tech. math. de la phys.
 OLIVIER Antropologie.
 NOZERAN T Bota ique II (Orsay).
 BLANC-LAPIERRE .. Physique E.N.S. (Orsay).
 CHEVAUGEON T Botanique (Orsay).
 KAHANE T Mathématiques (Orsay).
 SAMUEL T Math. E.N.S féminine.
 KOURGANOFF T Astronomie (Orsay).
 THIRY T Mécanique céleste.

Secrétaire Général R. POUILLAIN

PRÉVISION ET FILTRAGE LINÉAIRES
APPLICATION A LA REPRÉSENTATION CANONIQUE DES PROCESSUS
ALÉATOIRES DU SECOND ORDRE
ET A LA THÉORIE DE L'INFORMATION

par H. KÖREZLIOGLU,

Ingénieur

à la Société Télécommunications Radioélectriques et Téléphoniques

RÉSUMÉ

La première partie du présent travail est consacrée à la recherche des conditions dans lesquelles une fonction aléatoire du second ordre, non stationnaire, dont on connaît certaines propriétés simples telles que la fonction de covariance, peut être engendrée par le filtrage, variable dans le temps, d'un bruit blanc au sens large (accroissements d'un processus aléatoire à accroissements non corrélés). Ce problème se situe dans le cadre plus général de la représentation canonique des processus aléatoires, c'est-à-dire de leur représentation en fonction de leurs innovations (éléments non corrélés avec le passé des processus, antérieurs à l'instant où ils apparaissent au cours du déroulement des processus). Après une étude préliminaire du cas des processus aléatoires à temps discret, une solution est donnée pour une classe de fonctions aléatoires, qui contient aussi, comme cas particulier, les fonctions aléatoires stationnaires à densité spectrale rationnelle, et, pour cette classe de fonctions aléatoires, la réponse impulsionnelle du filtre est explicitée.

La deuxième partie de l'exposé est consacrée à l'étude du rôle joué dans la théorie de l'information par la prévision et le filtrage linéaires. Tenant compte de l'impossibilité pratique d'un délai infiniment grand pour l'estimation à la réception des signaux émis, on donne, en fonction de la valeur admissible de ce délai, l'expression du débit d'information dans le cas des canaux laplaciens, comme application de la théorie de la prévision et du filtrage linéaires. Et on retrouve, en augmentant indéfiniment ce délai, la formule de Shannon de la capacité d'un canal stationnaire, de bande de fréquence limitée, perturbé par un bruit blanc, de puissance moyenne connue, indépendant des signaux émis lorsque la puissance moyenne de ces signaux est connue.

PLAN

CHAPITRE I. — PRÉVISION LINÉAIRE ET PROBLÈME DE REPRÉSENTATION CANONIQUE DES PROCESSUS ALÉATOIRES DU SECOND ORDRE.

1.1. PROCESSUS DISCRETS DU SECOND ORDRE.

1.1.1. Représentation orthogonale canonique.

1.1.2. Une représentation intégrale des processus discrets du second ordre.

1.2. REPRÉSENTATION INTÉGRALE CANONIQUE DES PROCESSUS PERMANENTS DU SECOND ORDRE.

1.2.1. Préliminaires et définition d'une fonction d'innovation.

1.2.2. Conditions d'existence de cette fonction d'innovation.

1.2.3. Cas où la variance de cette fonction d'innovation est absolument continue.

1.2.4. Processus à projections analytiques.

1.2.5. Conclusions : justification des hypothèses ; une représentation intégrale duale ; cas des processus aléatoires stationnaires du second ordre.

CHAPITRE II. — PRÉVISION, FILTRAGE ET INFORMATION.

2.1. DÉBITS D'ENTROPIE ET D'INFORMATION.

2.2. PRÉVISION LINÉAIRE DES PROCESSUS VECTORIELS ET STATIONNAIRES DU SECOND ORDRE.

2.3. CALCUL DU DÉBIT D'ENTROPIE D'UN PROCESSUS LAPLACIEN STATIONNAIRE.

2.4. FILTRAGE LINÉAIRE DES PROCESSUS STATIONNAIRES ET DÉBIT D'INFORMATION TRANSMISE ENTRE DEUX PROCESSUS LAPLACIENS, STATIONNAIRES ET STATIONNAIREMENT CORRÉLÉS.

2.5. APPLICATIONS.

2.5.1. Un cas particulier concernant les processus laplaciens non stationnaires.

2.5.2. Confirmation de la formule de capacité de Shannon.

2.5.3. Transmission des innovations.

PRÉAMBULE.

Le présent travail est consacré à l'étude de la représentation canonique des processus aléatoires du second ordre et au calcul du débit d'entropie et d'information des processus aléatoires laplaciens, comme application de la théorie de la prévision et du filtrage linéaires.

Le problème de la représentation canonique des processus aléatoires du second ordre non stationnaires, a été introduit par P. Lévy (1951) sur l'exemple des processus aléatoires ayant pour représentation canonique

$$(1) \quad x(t) = \int_0^t c(t, s) d\xi(s),$$

où $\xi(t)$ est une fonction aléatoire à accroissements homogènes non-corrélés, c'est-à-dire telle que $E |d\xi(t)|^2 = dt$; et $c(t, s)$, une fonction certaine à carré du module sommable, sur $[0, t]$, par rapport à s . Parmi les processus aléatoires ayant une telle représentation, les propriétés de ceux qui vérifient l'équation aux accroissements du premier ordre

$$(2) \quad dx(t) = c(t, t) d\xi(t) + y(t) dt, \quad \text{avec}$$

$$(3) \quad y(t) = \int_0^t \frac{\partial}{\partial t} c(t, s) d\xi(s),$$

ont été étudiées par P. Lévy.

Mais le problème inverse de la recherche des conditions auxquelles doit satisfaire un processus pour qu'il admette une telle représentation restait non résolu. C'est ce problème qui constitue l'essentiel de notre travail sur la représentation canonique des processus aléatoires du second ordre.

En fait, le problème posé dans ses termes les plus généraux a trouvé sa réponse dans les travaux de Hida (1960) et de Cramer (1961) qui ont démontré l'existence d'une suite orthogonale $\{Z_n(t)\}$ de fonctions aléatoires à accroissements non corrélés telles que

$$(4) \quad x(t) = \sum_n \int_{-\infty}^t g_n(t, s) dZ_n(s).$$

Mais ni le calcul des *fonctions d'innovation* $Z_n(t)$, ni celui des noyaux $g_n(t, s)$ ne sont explicités dans leurs travaux.

Nous nous sommes donné, dans ce travail, le but de rechercher une classe de processus aléatoires du second ordre, dont la représentation canonique puisse être construite à partir de leurs propriétés temporelles simples.

Nous avons tout d'abord donné, dans le premier chapitre, en partant de la représentation canonique des processus à paramètre discret obtenue, par Cramer (1958), une représentation spectrale qui rappelle de très près la représentation harmonique des processus stationnaires, pour en déduire ensuite que tout processus du second ordre non stationnaire peut être subordonné à un processus stationnaire ; problème qui a son importance dans la reproduction des processus non stationnaires.

Mais cette analogie de représentation n'existe plus entre les processus permanents stationnaires et non stationnaires, comme le montre la représentation (4), et la technique de la prévision linéaire n'a plus un apport direct pour le calcul des fonctions d'innovation $Z_n(t)$ et des noyaux $g_n(t, s)$. Nous avons tout de même trouvé une classe de processus aléatoires pour laquelle nous avons pu expliciter une fonction d'innovation et une composante admettant une représentation canonique par la technique de la prévision linéaire. Cette classe est constituée par les processus tels que si $\mathbf{x}(t/s)$ désigne la prédiction linéaire de $x(t)$ en fonction des valeurs prises depuis $-\infty$ jusqu'à l'instant s , $\mathbf{x}(t/s)$ est une fonction aléatoire de t , dérivable en moyenne quadratique pour $t > s$, et nous appelons $x(t)$ processus à projections dérivables. Des conditions assez larges sur ces processus ont été données, pour qu'ils admettent une représentation du type (1) et qu'ils vérifient l'équation (2) de Lévy. Dans ces conditions, l'expression du noyau a été écrite sous forme explicite.

En conservant cette base de recherche, nous avons démontré qu'il existe une décomposition orthogonale

$$(5) \quad x(t) = u(t) + v(t),$$

où $u(t)$ est la partie purement non déterministe et $v(t)$, la partie déterministe du processus, sous l'hypothèse que $E[\mathbf{x}(t/s), \mathbf{x}(\tau/s)]$ est une fonction analytique de t, τ et de s ($s \leq t, \tau$). Mais en plus, notre démonstration s'appuyant sur un mode de construction du processus d'innovation, il nous a été possible de donner, sous l'hypothèse faite, l'expression du noyau canonique correspondant à $u(t)$ qui a pour représentation

$$(6) \quad u(t) = \int_{-\infty}^t c(t, s) d\xi(s).$$

Nous avons démontré aussi que $v(t)$ est un processus analytique.

L'hypothèse faite sur l'analyticité de $E[\mathbf{x}(t/s), \mathbf{x}(\tau/s)]$ pour t et τ supérieurs à s entraîne que $\mathbf{x}(t/s)$ est une fonction analytique de $(t-s)$. Pour cette raison, nous appelons les processus ayant cette propriété, processus à projections analytiques. Si l'on pense que $\mathbf{x}(t/s)$ considéré comme un processus aléatoire qui coïncide avec $x(t)$ jusqu'à l'instant s , est un processus sans innovation à partir de l'instant s , cette partie de l'hypothèse ne paraît pas être une restriction importante contrairement à l'analyticité de $E[\mathbf{x}(t/s), \mathbf{x}(\tau/s)]$ en s ($s \leq t, \tau$), qui assure une permanence aux innovations du processus. Nous suggérons, néanmoins, la méthode exposée comme une base de recherche des fonctions d'innovations des processus aléatoires.

Remarquons que tout processus stationnaire à densité spectrale rationnelle vérifie l'hypothèse faite ci-dessus et que la méthode utilisée est applicable, telle qu'elle est, au cas des processus définis seulement dans un intervalle fini.

Le second chapitre est consacré à l'étude de la liaison existant entre la prévision et le filtrage linéaires des processus aléatoires et l'information transmise entre deux processus vectoriels laplaciens. Nous avons tenu compte, dans l'expression du débit d'information, de l'impossibilité pratique, à la réception, d'un délai d'attente infiniment grand pour l'estimation des signaux émis. Nous avons donné, en utilisant une méthode due à Pinsker (1958), l'expression du débit d'information en fonction du retard admissible à la réception et montré que l'expression obtenue se ramène à celle donnée déjà par Pinsker lorsque le retard admissible est infiniment grand.

Cette dernière expression nous a permis de retrouver la formule de capacité de Shannon (1949), d'une ligne à bande de fréquence limitée à $[-W, W]$, perturbée par un bruit blanc de puissance moyenne B indépendant des signaux émis lorsque la puissance moyenne de ces signaux est limitée à une valeur donnée S .

CHAPITRE I : PRÉVISION LINÉAIRE ET PROBLÈMES DE REPRÉSENTATION CANONIQUE DES PROCESSUS ALÉATOIRES DU SECOND ORDRE.

Dans cette partie de notre étude nous allons considérer des processus aléatoires du second ordre à paramètre-temps discret ou continu. Nous supposons que toutes les *variables aléatoires* (*v.a.*) constituant le processus considéré possèdent une espérance mathématique nulle. Pour simplifier le langage et l'écriture, nous utiliserons aussi les expressions *suite aléatoire* pour les processus à paramètre discret et *fonction aléatoire* pour les processus à paramètre continu et nous désignerons par le même symbole les v.a. et les valeurs prises par elles.

Dans l'espace de Hilbert $\mathcal{E}(x)$ engendré par une suite aléatoire $\{x_n\}$ (n prenant toutes les valeurs entières) ou par une fonction aléatoire $x(t)$, (t prenant toutes les valeurs réelles), la norme d'un élément u , définie par $(E|u|^2)^{1/2}$ sera désignée par $\|u\|$, et le produit scalaire de deux éléments u et v , défini par $E(u \cdot v)$ sera désigné par $\langle u, v \rangle$. (E : espérance mathématique).

1.1. PROCESSUS DISCRETS DU SECOND ORDRE.

1.1.1. Représentation orthogonale canonique.

Dans le but de reprendre certaines définitions utilisées, dans la théorie de la prévision linéaire, nous allons, tout d'abord, donner sans démonstration, le théorème de décomposition orthogonale des processus discrets du second ordre, démontré par H. Cramer [1] et qui est une extension du théorème bien connu dans le cas des processus stationnaires (cf. par exemple : Grenander & Rosenblatt, p. 70).

Etant donnée une suite aléatoire $\{x_n\}$ du second ordre, il est possible de faire correspondre à l'élément x_{n+p} du sous-espace linéaire fermé $\mathcal{E}\{x_j : n-q \leq j \leq n+p\}$ de $\mathcal{E}(x)$, un élément y_q de $\mathcal{E}\{x_j : n-q \leq j \leq n\}$ tel que $\|x_{n+p} - z_q\|$ soit minimum pour $z_q = y_q$, pour tout

$$z_q \in \mathcal{E}\{x_j : n-q \leq j \leq n\}.$$

Nous avons alors la relation d'orthogonalité suivante :

$$x_{n+p} - y_q \perp \mathcal{E}\{x_j : n-q \leq j \leq n\}.$$

y_q s'appelle la projection de x_{n+p} sur $\mathcal{E}\{x_j : n-q \leq j \leq n\}$ et $x_{n+p} - y_q$ l'erreur de prévision sur x_{n+p} , connaissant $x_{n-q}, \dots, x_n$.

On démontre que, lorsque $q \rightarrow \infty$, la quantité

$$\|x_{n+p} - y_q\| = \alpha_q$$

tend vers une valeur $\alpha \leq \alpha_q$ et qu'il existe un élément y de $\mathcal{E}\{x_j : j \leq n\}$ tel que

$$\|x_{n+p} - y\| = \alpha \quad \text{et que} \quad \lim_{q \rightarrow \infty} \|y - y_q\| = 0.$$

On dit que le processus x_n est *déterministe* si $\alpha = 0$, quels que soient $p \leq 1$ et n .

Soit $\mathbf{x}_n$ l'élément de $\mathcal{E}_{n-1}(x) = \mathcal{E}\{x_j : j \leq n-1\}$, défini par

$$\|x_n - \mathbf{x}_n\| = \min_{z \in \mathcal{E}_{n-1}(x)} \|x_n - z\|$$

qui, par conséquent, vérifie la relation

$$x_n - \mathbf{x}_n \perp \mathcal{E}_{n-1}(x).$$

Posons

$$(1.1.1) \quad x_n - \mathbf{x}_n = c_{n,n} \xi_n \quad \text{avec} \quad c_{n,n} = \|x_n - \mathbf{x}_n\|.$$

On dit que $x_n - \mathbf{x}_n$ est l'*innovation* du processus $\{x_n\}$ à l'instant n . La v. a. ξ_n est bien définie toutes les fois que $\|x_n - \mathbf{x}_n\| \neq 0$. On appelle *indices d'indétermination* du processus $\{x_n\}$, l'ensemble des valeurs de n , pour lesquelles $\|x_n - \mathbf{x}_n\| \neq 0$. Soit $\{j_k\}$ la suite des indices d'indétermination du processus $\{x_n\}$. La suite $\{\xi_{j_k}\}$, définie par (1.1.1) pour $j_k \leq n$, constitue une base orthonormale du sous-espace $\mathcal{E}_n(X)$. Alors l'élément x_n admet *presque sûrement* (*p. s.*) le développement orthogonal suivant :

$$(1.1.2) \quad x_n = \sum_{j_k \leq n} c_{n,j_k} \xi_{j_k} + v_n, \quad \text{avec} \quad c_{n,j_k} = \langle x_n, \xi_{j_k} \rangle$$

où la v. a. v_n est orthogonale à tous les ξ_{j_k} .

Considérons la suite

$$(1.1.3) \quad u_n = \sum_{j_k \leq n} c_{n,j_k} \xi_{j_k}.$$

Tout élément de cette suite ayant une variance finie, d'après

$$\|u_n\|^2 = \sum_{j_k \leq n} |c_{n,j_k}|^2 \leq |x_n|^2 \leq \infty,$$

nous voyons que $c_{n,j_k} \rightarrow 0$, lorsque $j_k \rightarrow -\infty$. Cela veut dire que l'élément u_n est *non corrélé avec le passé infiniment lointain du processus* $\{x_n\}$ ou bien que l'espace

$$\bigcap_{q=0}^{\infty} \mathcal{E}_{n-q}(u)$$

est vide. Le processus $\{u_n\}$ possède les mêmes innovations que le processus $\{x_n\}$. Si u_n désigne la projection de u_n sur $\mathcal{E}_{n-1}(u)$, nous avons

$$(1.1.4) \quad \|u_n - u_n\| = \|x_n - x_n\| = c_{n,n}; \quad u_n - u_n = c_{n,n} \xi_n.$$

Ceci implique que le processus $\{v_n\}$ est un processus déterministe. Pour un tel processus, les sous-espaces $\mathcal{E}_n(v)$ sont les mêmes que soit n .

DÉFINITIONS 1.1.1. — 1° Etant donnée une suite $\{\xi_{j_k}\}$ de v. a. orthonormales, nous dirons que la représentation orthogonale

$$x_n = \sum_{j_k \leq n} c_{n,j_k} \xi_{j_k},$$

du processus $\{x_n\}$ est une représentation canonique, si

$$\sum_{j_k \leq n-p} c_{n,j_k} \xi_{j_k}$$

est la projection de x_n sur le sous-espace $\mathcal{E}_{n-p}(x)$ quels que soient $p \geq 0$ et n . Il en est, par exemple, ainsi pour la représentation (1.1.3) du processus $\{u_n\}$.

2° Nous dirons qu'un processus non-corrélé avec son passé infiniment lointain est un processus purement non déterministe, si la suite des indices d'indétermination coïncide avec l'ensemble de tous les entiers. En particulier, c'est le cas des processus stationnaires du second ordre, non-corrélés avec leurs passés infiniment lointains.

1.1.2. Une représentation intégrale des processus discrets du second ordre.

Considérons une suite de v. a. orthonormale $\{\xi_n\}$, définie pour toutes les valeurs entières de l'indice n ; c'est-à-dire telle que

$$(1.2.1) \quad \langle \xi_n, \xi_m \rangle = \begin{cases} 1, & \text{pour } n = m, \\ 0, & \text{pour } n \neq m. \end{cases}$$

Les coefficients de corrélation (1.2.1) de cette suite peuvent aussi être donnés par l'intégrale suivante :

$$(1.2.2) \quad \langle \xi_n, \xi_m \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)\lambda} d\lambda = \begin{cases} 1, & \text{pour } n = m, \\ 0 & \text{pour } n \neq m. \end{cases}$$

Selon cette représentation, un théorème dû à Karhunen ([1], satz 10) entraîne le résultat suivant :

LEMME 1.2.1. — Etant donnée une suite $\{\xi_n\}$ de v. a. orthonormales définies pour toutes les valeurs entières de n , il lui correspond une fonction aléatoire $\xi(\lambda)$ à accroissements homogènes non corrélés, uniquement définie dans $[-\pi, \pi]$, telle que

$$(1.2.3) \quad \xi_n = \int_{-\pi}^{\pi} e^{in\lambda} d\xi(\lambda) \quad \text{avec } \|d\xi(\lambda)\|^2 = d\lambda/2\pi.$$

On peut définir (cf. Cramer [2]), à partir d'une telle fonction $\xi(\lambda)$ une fonction aléatoire d'ensemble qui prend la valeur

$$\xi(S) = \int_S d\xi(\lambda)$$

sur tout ensemble borélien S de l'intervalle $[-\pi, +\pi]$. Désignons par $\mathcal{E}(\xi)$ la fermeture de l'espace linéaire constitué par des éléments de la forme $a_1 \xi(S_1) + a_2 \xi(S_2) + \dots + a_n \xi(S_n)$ où les a_j sont des coefficients complexes et S_j des ensembles boréliens de $[-\pi, \pi]$. L'élément défini par l'intégrale stochastique

$$(1.2.4) \quad x = \int_{-\pi}^{\pi} \gamma_x(\lambda) d\xi(\lambda)$$

où $\gamma_x(\lambda)$ est une fonction mesurable et telle que l'intégrale

$$(1.2.5) \quad \int_{-\pi}^{\pi} |\gamma_x(\lambda)|^2 d\lambda$$

existe, est un élément de $\mathcal{E}(\xi)$. Le produit scalaire de deux éléments x et y ainsi défini est

$$(1.2.6) \quad \langle x, y \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \gamma_x(\lambda) \cdot \overline{\gamma_y(\lambda)} d\lambda$$

L'espace linéaire $L_2(-\pi, \pi)$ de toutes les fonctions $\gamma(\lambda)$ mesurables, définies dans $[-\pi, \pi]$ et pour lesquelles l'intégrale (1.2.5) existe, est un espace de Hilbert, lorsque dans cet espace, le produit scalaire de deux fonctions γ_x et γ_y est défini par l'intégrale (1.2.6). L'intégrale (1.2.4) constitue donc une transformation biunivoque, linéaire et isométrique entre les deux espaces $\mathcal{E}(\xi)$ et $L_2(-\pi, \pi)$. Par conséquent, à toute suite de fonctions orthonormales dans $L_2(-\pi, \pi)$ correspond une suite de v. a. orthonormales dans $\mathcal{E}(\xi)$. La relation (1.2.3) n'est autre qu'une telle correspondance. Nous voyons alors que l'espace de Hilbert engendré par la suite $\{\xi_n\}$ coïncide avec $\mathcal{E}(\xi)$.

THÉOREME 1.2.1. — *Etant donné un processus $\{u_n\}$, non-corrélé avec son passé infiniment lointain, il lui correspond au moins une fonction aléatoire $\xi(\lambda)$ à accroissements homogènes, non-corrélés, définie dans $[-\pi, \pi]$, telle que :*

$$(1.2.7) \quad u_n = \int_{-\pi}^{\pi} e^{in\lambda} \gamma_n(\lambda) d\xi(\lambda), \quad \|d\xi(\lambda)\|^2 = \frac{d\lambda}{2\pi},$$

où $\gamma_n(\lambda)$ est une fonction à carré du module sommable et admet un développement de Fourier

$$(1.2.8) \quad \sum_{j_k \leq n} c_{n,j_k} e^{i(j_k - n)\lambda},$$

qui converge en norme dans $L_2(-\pi, \pi)$, les coefficients c_{n,j_k} étant définis par (1.1.3).

Remarquons que, dans le cas des processus stationnaires du second ordre, la représentation (1.2.7) est connue ; auquel cas, la fonction (1.2.8) qui s'appelle fonction génératrice du processus $\{u_n\}$, ne dépend pas du temps et peut se mettre sous la forme

$$\gamma(\lambda) = \sum_{m=0}^{\infty} c_m e^{-im\lambda} \quad (\text{cf. par exemple Doob}).$$

Démonstration. — Considérons, tout d'abord, un processus purement non déterministe. Nous pouvons définir pour toutes les valeurs entières de n , une suite de v. a. orthonormales, par

$$\xi = \frac{u_n - u_n}{\|u_n - u_n\|}$$

Nous avons vu que u_n admet, p. s., le développement orthogonal

$$u_n = \sum_{m=0}^{\infty} c_{n,n-m} \xi_{n-m}, \quad \text{où } c_{n,n-m} = \langle u_n, \xi_{n-m} \rangle.$$

La connaissance de la suite $\{\xi_n\}$ étant alors suffisante pour la détermination de la fonction spectrale $\xi(\lambda)$, d'après le lemme 1.2.1, nous voyons que u_n peut être représenté par

$$(1.2.9) \quad u_n = \lim_{M \rightarrow \infty} \text{m.q.} \int_{-\pi}^{\pi} e^{in\lambda} \left[\sum_{m=0}^M c_{n,n-m} e^{-im\lambda} \right] d\xi(\lambda) = \int_{-\pi}^{\pi} e^{in\lambda} \gamma_n(\lambda) d\xi(\lambda)$$

$$\text{puisque } \sum_{m=0}^M c_{n,n-m} e^{-im\lambda} \text{ converge en norme dans } L_2(-\pi, \pi).$$

Mais dans le cas où les indices d'indétermination du processus ne prennent pas toutes les valeurs entières, la connaissance du processus ne suffit plus à déterminer la fonction spectrale $\xi(\lambda)$. Il est aisé de voir qu'il est toujours possible de trouver une fonction spectrale, telle que le processus $\{u_n\}$ admette la représentation (1.2.7) ; mais cette fois-ci, sa détermination n'est pas unique. En effet, il suffit de compléter la suite $\{\xi_{j_k}\}$ par une suite orthonormale $\{\gamma_{j_l}\}$ orthogonale à $\{\xi_{j_k}\}$ et définie sur l'ensemble complémentaire de l'ensemble des indices d'indétermination et de définir le processus $\{u_n\}$ dans l'espace de Hilbert engendré par la suite

$$\xi_n = \begin{cases} \xi_{j_k} & \text{pour } n = j_k \\ \gamma_{j_l} & \text{pour } n = j_l \end{cases}$$

pour, finalement, retrouver la représentation (1.2.7).

THÉORÈME 1.2.2. — La fonction

$$(1.2.10) \quad \Gamma_n(z) = \sum_{j_k \leq n} \bar{c}_{n,j_k} z^{n-j_k}$$

possède les propriétés suivantes : 1°) elle est analytique pour $|z| < 1$;

$$(1.2.11) \quad 2^\circ) \lim_{r \nearrow 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} |\Gamma_n(r e^{i\lambda})|^2 d\lambda = \sum_{j_k \leq n} |c_{n,j_k}|^2 ;$$

$$3^\circ) \lim_{r \nearrow 1} \bar{\Gamma}_n(r e^{i\lambda}) = \gamma_n(\lambda) \text{ pour presque tout } \lambda.$$

$$4^\circ) |\gamma_n(\lambda)| \text{ est intégrale au sens de Lebesgue et}$$

$$(1.2.12) \quad \log |c_{n,n}|^2 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |\gamma_n(\lambda)|^2 d\lambda.$$

Dans le cas des processus stationnaires du second ordre purement non déterministes, $\Gamma_n(z)$ qui ne dépend pas de n , n'a pas de zéros pour $|z| < 1$ et seule l'égalité est valable dans (1.2.12).

Démonstration. — 1° En posant

$$(1.2.13) \quad \Gamma_n(z) = \sum_{j_k \leq n} \bar{c}_{n,j_k} r^{n-j_k} e^{i(n-j_k)\lambda}, \quad \text{pour } z = r e^{i\lambda}$$

$$\text{nous avons, d'une part, } |\Gamma_n(z)|^2 = \sum_{p=0}^{\infty} r^p e^{-ip\lambda} \left[\sum_{n-p \leq j_k \leq n} \bar{c}_{n,j_k} \cdot c_{n,2n-j_k-p} e^{2(n-j_k)\lambda} \right] ;$$

$$\text{et, d'autre part, la relation } c_{n,j_k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(j_k-n)\lambda} \gamma(\lambda) d\lambda \quad \text{entraîne}$$

$$|c_{n,j_k} c_{n,2n-j_k-p}| \leq \frac{1}{4\pi^2} \left(\int_{-\pi}^{\pi} \gamma(\lambda) d\lambda \right)^2 = A < \infty. \quad \text{Donc}$$

$$|\Gamma_n(z)|^2 \leq \sum_{p=0}^{\infty} r^p \left[\sum_{n-p \leq j_k \leq n} |\bar{c}_{n,j_k} \cdot c_{n,2n-j_k-p}| \right] \leq A \sum_{p=0}^{\infty} (p+1) r^p = A \cdot \frac{1}{(r-1)^2} < \infty, \text{ pour } r < 1.$$

2° Nous avons, pour $r < 1$,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |\Gamma_n(r e^{i\lambda})|^2 d\lambda = \sum_{j_k \leq n} r^{2(n-j_k)} |c_{n,j_k}|^2 < \sum_{j_k \leq n} |c_{n,j_k}|^2 < \infty.$$

La série $\sum_{j_k \leq n} |c_{n,j_k}|^2 (1 - r^{2(n-j_k)})$, converge uniformément dans $[0, 1]$. Puisque toute somme finie

$$-N \leq \sum_{j_k \leq n} |c_{n,j_k}|^2 (1 - r^{2(n-j_k)})$$

est continue en r dans $[0, 1]$, il en est de même pour la limite lorsque $N \rightarrow \infty$; d'où la propriété du 2°.

Grâce à cette propriété que nous venons de démontrer, la fonction (1.2.13) appartient à la classe H_2 de Hardy (cf. Grenander et Rosenblatt, appendix ou complex variable theory).

3° Comme $\gamma_n(\lambda) \in L_2(-\pi, \pi)$, $\gamma_n(\lambda) \in L_1(-\pi, \pi)$ aussi. Alors on démontre par l'intermédiaire de la formule de Nevanlinna (cf. référence ci-dessus) que

$$\lim_{r \nearrow 1} \bar{\Gamma}_n(r e^{i\lambda}) = \gamma_n(\lambda) \quad \text{pour presque tout } \lambda.$$

4° L'intégralité de $\log |\gamma_n(\lambda)|$ au sens de Lebesgue devient une conséquence de la formule de Jensen (cf. Grenander et Szegő, 1.13 (b)), ainsi que la relation (1.2.12) (cf. Valiron, p. 423), où l'égalité n'est valable que si $\Gamma_n(z)$ n'a pas de zéros pour $|z| < 1$. Nous allons maintenant montrer qu'il en est effectivement ainsi dans le cas des processus stationnaires du second ordre.

Soit

$$\Gamma_n(z) = \Gamma(z) = \sum_{m=0}^{\infty} \bar{c}_m z^m \quad \text{pour tout } n.$$

Supposons que cette fonction ait un zéro z_0 , ($|z_0| < 1$). Il est clair que $z_0 \neq 0$, sinon nous aurions $\Gamma(0) = c_0 = 0$; ce qui est contraire au fait que le processus $\{u_n\}$ est non déterministe. Considérons alors, le processus défini par

$$(1.2.14) \quad y_n = \sum_{m=0}^{\infty} z_0^m \xi_{n-m}.$$

C'est un processus autorégressif et nous pouvons écrire

$$(1.2.15) \quad \begin{cases} a) & y_n = z_0 y_{n-1} + \xi_n, \\ b) & y_n = z_0^2 y_{n-2} + z_0 \xi_{n-1} + \xi_n, \\ c) & y_n = z_0 y_{n-3} + z_0^2 \xi_{n-2} + z_0 \xi_{n-1} + \xi_n. \end{cases}$$

D'une part, nous voyons que y_n est un élément du sous-espace linéaire $\mathcal{E}_n(u)$ d'après (1.2.14). Mais si, d'autre part z_0 était un zéro de $\Gamma(z)$, nous aurions

$$\langle y_n, u_n \rangle = \langle y_{n-1}, u_{n-1} \rangle = \dots = \sum_{m=0}^{\infty} \bar{c}_m z_0^m = 0.$$

Alors $y_n \perp u_n$, mais aussi $y_n \perp u_{n-1}$ d'après a); puisque $y_{n-1} \perp u_{n-1}$ et $\xi_n \perp u_{n-1}$. On pourrait aussi voir que $y_n \perp u_{n-2}$, d'après b). Nous pourrions montrer de cette façon que $y_n \perp \mathcal{E}_n(u)$. Ce qui est contradictoire avec le fait que y_n est un élément de $\mathcal{E}_n(u)$. Donc z_0 ne peut pas être un zéro de $\Gamma(z)$.

DÉFINITION 1.2.1. — Une suite aléatoire $\{x_n\}$ est dite *subordonnée* à une autre suite aléatoire $\{y_n\}$, si, pour toute valeur de n ,

$$\mathcal{E}_n(x) \subseteq \mathcal{E}_n(y).$$

Les suites $\{x_n\}$ et $\{y_n\}$ sont dites *coordonnées*, si

$$\mathcal{E}_n(x) = \mathcal{E}_n(y).$$

(Cette définition est légèrement plus restrictive que celle donnée par Kolmogorov et Karhunen [2], en ce sens qu'ici nous considérons les sous-espaces engendrés par les éléments du processus jusqu'à l'instant n et non pas l'espace de Hilbert engendré par tout le processus).

Remarquons que le théorème 1.2.1 revient à dire que, pour tout processus $\{u_n\}$ non-corrélé avec son passé infiniment éloigné, on peut trouver un processus orthonormal $\{\xi_n\}$ auquel il est subordonné.

Nous avons vu que les sous-espaces $\mathcal{E}_n(v)$ engendrés par une suite déterministe $\{v_n\}$ sont tous les mêmes quel que soit n . L'espace $\mathcal{E}(v)$ étant séparable, désignons par $\{\eta_k\}$ le système de v. a. orthonormales qui constitue une base complète dans $\mathcal{E}(v)$. Cette suite peut être finie ou infinie selon les cas. Définissons une suite d'éléments $\{w_n\}$ dans $\mathcal{E}(v)$ par

$$w_n = \sum_k a_k e^{in\lambda_k} \eta_k, \quad \text{avec } -\pi < \lambda_k \leq \pi$$

telle que sa connaissance puisse déterminer uniquement la suite $\{\eta_k\}$ (il est toujours possible de définir une telle suite); $\{w_n\}$ constitue un processus stationnaire; car, le produit scalaire

$$\langle w_m, w_n \rangle = \sum_k |a_k|^2 e^{i(m-n)\lambda_k}$$

ne dépend que de la différence $(m - n)$. Cette suite est aussi une suite déterministe. En effet, considérons une fonction en escalier $F_w(\lambda)$, définie dans $[-\pi, +\pi]$ dont les sauts prennent les valeurs $|a_k|^2$ aux points λ_k . Cette fonction peut être considérée comme la fonction de répartition spectrale du processus $\{w_n\}$ qui admet, alors, la représentation intégrale

$$w_n = \int_{-\pi}^{\pi} e^{in\lambda} dz_w(\lambda), \quad \|z_w(\lambda)\|^2 = F_w(\lambda),$$

$z_w(\lambda)$ étant une fonction aléatoire à accroissements non-corrélés, qui prend des accroissements $a_k \eta_k$ aux points λ_k et des accroissements p. s. nuls, partout ailleurs, dans $[-\pi, \pi]$. Il est connu (cf. Kolmogorov) qu'un tel processus est déterministe. Les accroissements de $z_w(\lambda)$ qui coïncident avec la base de $\mathcal{E}(v)$, constituent une base complète dans $\mathcal{E}(w)$.

Comme $\{v_n\}$ est une suite subordonnée à $\{w_n\}$, elle peut être représentée par

$$(1.2.16) \quad v_n = \int_{-\pi}^{\pi} e^{in\lambda} \delta_n(\lambda) dz_w(\lambda).$$

où $\delta_n(\lambda)$ est une fonction certaine à carré du module sommable par rapport à $F_w(\lambda)$.

Nous prendrons, comme il a été précisé plus haut, $a_k \lambda_k = z_w(\lambda_k + 0) - z_w(\lambda_k)$. D'après la définition de l'intégrale de Stieltjes-Lebesgue (cf. Doob)

$$(1.2.17) \quad v_n = \sum_k e^{in\lambda_k} \delta_n(\lambda_k) a_k \eta_k, \quad \text{avec}$$

$$(1.2.18) \quad \|v_n\|^2 = \int_{-\pi}^{\pi} |\delta_n(\lambda)|^2 dF_w(\lambda) = \sum_k |\delta_n(\lambda_k)|^2 |a_k|^2.$$

Compte tenu des représentations intégrales (1.2.7) et (1.2.16) tout processus aléatoire $\{x_n\}$ du second ordre peut se mettre sous la forme

$$(1.2.19) \quad x_n = \int_{-\pi}^{\pi} e^{in\lambda} \gamma(\lambda) d\xi(\lambda) + \int_{-\pi}^{\pi} e^{in\lambda} \delta_n(\lambda) dz_w(\lambda).$$

La fonction non décroissante définie par

$$(1.2.20) \quad F(\lambda) = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\mu + F_w(\lambda)$$

peut être considérée comme la fonction de répartition spectrale d'un processus $\{y_n\}$ stationnaire

$$(1.2.21) \quad y_n = \int_{-\pi}^{\pi} e^{in\lambda} dz(\lambda), \quad \|z(\lambda)\|^2 = F(\lambda).$$

La suite de v. a.

$$(1.2.22) \quad x_n^* = \int_{-\pi}^{\pi} e^{in\lambda} \gamma_n^*(\lambda) dz(\lambda),$$

où la fonction $\gamma_n^*(\lambda)$ ne diffère de $\gamma_n(\lambda)$ que sur l'ensemble $\{\lambda_k\}$ défini plus haut, coïncide p. s. avec la suite $\{x_n\}$ (1.2.19), lorsque $\gamma_n^*(\lambda)$ prend les valeurs $\delta_n(\lambda_k)$ sur cet ensemble. Car, nous avons

$$\gamma_n^*(\lambda) = \gamma_n(\lambda) \text{ presque partout (mesure de Lebesgue)}$$

$$\gamma_n^*(\lambda) = \delta_n(\lambda) \text{ presque partout (dF}_w(\lambda) \text{ — mesure)} \quad \text{et}$$

$$x_n^* - x_n = \int_{-\pi}^{\pi} e^{in\lambda} [\gamma_n^*(\lambda) - \gamma_n(\lambda)] d\xi(\lambda) + \int_{-\pi}^{\pi} e^{in\lambda} [\gamma_n^*(\lambda) - \delta_n(\lambda)] dz_w(\lambda)$$

Par conséquent, $\|x_n^* - x_n\| = \text{donc } 0$; $x_n^* = x_n$ p. s. Nous pouvons alors énoncer le théorème suivant :

THÉORÈME 1.2.3. — *Toute suite aléatoire du second ordre est subordonnée à une suite stationnaire dans le sens de la définition 1.2.1.*

1.2. REPRÉSENTATION INTÉGRALE CANONIQUE DES PROCESSUS PERMANENTS DU SECOND ORDRE.

1.2.1. Préliminaires et définition d'une fonction d'innovation.

Soit $x(t)$ un processus aléatoire du second ordre défini pour toutes les valeurs du paramètre-temps t , dans $(-\infty, +\infty)$ et désignons par $x(t/t_0)$, la projection de $x(t)$ sur le sous-espace linéaire $\mathcal{E}_{t_0} = \mathcal{E}\{x(\tau); \tau \leq t_0\}$ ($t_0 < t$).

Considérons un intervalle fini $[t_0, t]$ et soient

$$(2.1.1) \quad \{t_{N_k}\}_{k=0}^N = \{t_{N_0} = t_0, t_{N_1}, \dots, t_{N_N} = t\}$$

une subdivision quelconque de cet intervalle, et $S_N(t_0, t)$ la v. a. définie par

$$(2.1.2) \quad S_N(t_0, t) = \sum_{k=1}^N [x(t_{N_k}) - x(t_{N_k}/t_{N_{k-1}})].$$

Remarquons que tous les termes de cette série sont deux à deux orthogonaux, d'après la propriété suivante de la projection

$$(2.1.3) \quad x(t) - x(t/t_0) \perp \mathcal{E}_{t_0}.$$

Supposons que la somme $S_N(t_0, t)$ définie pour tout intervalle fini, tende en m. q. vers une fonction aléatoire $Z(t_0, t)$ uniquement déterminée, quelle que soit la subdivision $\{t_{N_k}\}$ lorsque $N \rightarrow \infty$ et que $\max(t_{N_k} - t_{N_{k-1}}) \rightarrow 0$. [Toutes les fois qu'il s'agira de la limite d'une somme dépendant d'une telle subdivision, il sera sous-entendu que $\max(t_{N_k} - t_{N_{k-1}}) \rightarrow 0$, lorsque $N \rightarrow \infty$]. Il est évident que la relation (2.1.3) entraîne, pour $t_0 < t_1 < t$,

$$Z(t_0, t_1) \perp Z(t_1, t).$$

De ce fait, $Z(t_0, t)$ qui est une fonction additive d'intervalle (cf. Cramer [2]) est à accroissements non corrélés en tant que fonction de t . Comme une telle fonction est définie à une constante additive près, nous poserons

$$(2.1.4) \quad Z(t_0, t) = Z_1(t) - Z_1(t_0), \text{ en prenant } Z_1(0) = 0.$$

Soit $\mathcal{E}_t(dZ_1)$ (cf. Cramer [2] et Karhunen [1]), l'espace de Hilbert des v. a. définies par

$$(2.1.5) \quad \int_{-\infty}^t g(s) dZ_1(s)$$

où les fonctions certaines $g(s)$ sont mesurables et sont telles que l'intégrale

$$(2.1.6) \quad \int_{-\infty}^t |g(s)|^2 dF_1(s), \quad \text{où } F_1(s) = \|Z_1(s)\|^2$$

existe; cette dernière intégrale définissant la norme de l'élément (2.1.5). Nous avons évidemment :

$$(2.1.7) \quad \mathcal{E}_t(dZ_1) \subseteq \mathcal{E}_t(x).$$

Désignons par $x_1(t)$ la projection de $x(t)$ sur $\mathcal{E}_t(dZ_1)$; $x_1(t)$ peut être représenté par :

$$(2.1.8) \quad x_1(t) = \int_{-\infty}^t g_1(t, s) dZ_1(s).$$

Nous nous attacherons dans ce qui suit, à l'étude de l'existence de la fonction $Z_1(t)$ et aux propriétés de $x(t) - x_1(t)$, dans le but de donner une représentation du processus $x(t)$, liée à ses innovations. Nous allons d'abord, donner quelques définitions, pour délimiter le cadre de notre étude.

DÉFINITIONS 2.1.1. — 1° (P. Lévy [3]). — Une fonction aléatoire $x(t)$ admet une représentation canonique, s'il existe une fonction aléatoire $Z(t)$ à accroissements non-corrélés, et une fonction certaine $c(t, s)$, telles que

$$x(t) = \int_{-\infty}^t c(t, s) dZ(s) \quad \text{p. s.}$$

$$\|x(t)\|^2 = \int_{-\infty}^t |c(t, s)|^2 dF(s), \quad F(s) = \|Z(s)\|^2, \quad x(t/t_0) = \int_{-\infty}^t c(t, s) dZ(s) \quad (\text{pour } t_0 < t).$$

2° (P. Lévy [3]). — Une représentation canonique est dite propre, si

$$\mathcal{E}_t(x) = \mathcal{E}_t(dZ).$$

P. Lévy a fait une étude détaillée (cf. Lévy [1], [2], [3]) des processus qui admettent une représentation canonique de la forme

$$x(t) = \int_0^t c(t, s) d\xi(s)$$

où $\xi(s)$ est un processus à accroissements homogènes non corrélés, c'est-à-dire tel que $\|d\xi(s)\|^2 = ds$. Mais le problème inverse posé par lui, à savoir, le problème de trouver les conditions auxquelles doit satisfaire un processus aléatoire pour qu'il admette une représentation canonique, a trouvé sa réponse dans le travail de T. Hida, basé sur les propriétés spectrales de l'opérateur de projection dans l'espace de Hilbert dont la fonction de covariance du processus considéré est le noyau reproduisant, et dans

le travail de Cramer [3], basé sur les propriétés structurales de l'espace de Hilbert engendré par le processus considéré. Nous nous contenterons de donner ici, à titre de comparaison avec le présent travail, l'énoncé du théorème sur la représentation canonique généralisée des processus, dû à Hida et à Cramer indépendamment.

THÉORÈME 2.1.1. — *Etant donné $x(t)$, un processus du second ordre, si $\mathcal{E}_t(x)$ est séparable quel que soit t , et si l'espace limite $\lim_{t \rightarrow -\infty} \mathcal{E}_t(x)$ est vide, alors, il existe une suite orthogonale, dénombrable, de fonctions aléatoires $Z_n(t)$ à accroissement non corrélés, telles que*

$$(2.1.9) \quad x(t) = \sum_n \int_{-\infty}^t g_n(t, s) dZ_n(s) \quad (\text{p. s.}) \quad \text{avec}$$

$$\|x(t)\|^2 = \sum_n \int_{-\infty}^t |g_n(t, s)|^2 dF_n(s), \quad \text{où } F_n(s) = \|Z_n(s)\|^2 \quad \text{et, que pour } t_0 < t$$

$$(2.1.10) \quad x(t/t_0) = \sum_n \int_{-\infty}^{t_0} g_n(t, s) dZ_n(s).$$

On voit que la représentation (2.1.9) sera une représentation canonique simple au sens de Lévy, si la suite $\{Z_n(t)\}$ est réduite à une seule fonction $Z(t)$.

Nous nous intéresserons ici au cas régulier, où toutes les fonctions $Z_n(t)$ seront des fonctions à accroissements homogènes non corrélés, qui, toutes, seront déduites à partir des fonctions aléatoires d'ensemble définies de la même manière que la fonction $Z_1(t)$ introduite plus haut.

1.2.2. Conditions d'existence de la fonction d'innovation $Z_1(t)$.

DÉFINITION 2.2.1. — Nous dirons que $x(t)$ est à projections dérivables (en m. q.) si, $x(t/t_0)$ est dérivable, en m. q., comme fonction de t , quels que soient t et t_0 ($t < t_0$).

LEMME 2.2.1. — La condition nécessaire et suffisante pour qu'un processus $x(t)$ soit à projections dérivables en m. q. est que la dérivée seconde généralisée de la fonction de covariance $\Gamma_{t_0}(t, \tau)$ de $x(t/t_0)$ existe et soit finie pour tous les points de la diagonale $t = \tau > t_0$.

Ce lemme n'est qu'une version du théorème déjà connu (cf. Loeve) sur la dérivabilité en m. q. pour le processus $x(t/t_0)$ défini dans (t_0, ∞) .

Remarque. — Nous supposerons, pour ce qui suit, que $x(t)$ est continu en m. q. et que les limites

$$(2.2.1) \quad \lim_{h \downarrow 0} \frac{1}{h} [x(t_0 + h/t_0) - x(t_0)], \quad \lim_{h \downarrow 0} \frac{1}{h} [x(t_0/t_0 - h) - x(t_0 - h)]$$

existent en m. q. et qu'elles sont égales. Nous désignerons la limite commune par $x'(t_0/t_0)$. L'hypothèse de la continuité en m. q. de $x(t)$ entraîne simplement que

$$(2.2.2) \quad \lim_{h \downarrow 0} (\text{m. q.}) x(t_0 + h/t_0) = \lim_{h \downarrow 0} (\text{m. q.}) x(t_0/t_0 - h) = x(t_0).$$

Quand nous parlerons des processus à projections dérivables, il sera sous-entendu que ces deux hypothèses sont satisfaites. Nous verrons, d'ailleurs, une généralisation de ces définitions dans le paragraphe suivant.

THÉORÈME 2.2.1. — *Soit $\Gamma(t, \tau)$, la fonction de covariance de $x(t)$, processus continu en m. q. et à projections dérivables. Si $x'(t/t)$ est continu en m. q., alors la somme*

$$(2.2.3) \quad S_N(t_0, t) = \sum_{k=1}^N [x(t_{Nk}) - x(t_{Nk}/t_{Nk-1})]$$

tend en m. q. vers une fonction aléatoire $Z(t_0, t)$, à accroissements en t non-corrélés, uniquement définie par

$$(2.2.4) \quad Z(t_0, t) = x(t) - x(t_0) - \int_{t_0}^t x'(s/s) ds.$$

et dont la norme est donnée par la limite

$$(2.2.5) \quad \|Z(t_0, t)\|^2 = F(t_0, t) = \lim_{N \rightarrow \infty} \sum_{k=1}^N [\Gamma(t_{Nk}, t_{Nk}) - \Gamma(t_{Nk}, t_{Nk-1}) - \Gamma(t_{Nk-1}, t_{Nk}) + \Gamma(t_{Nk-1}, t_{Nk-1})]$$

Démonstration. — Remarquons que $F(t_0, t)$ est la variation totale de la covariance $\Gamma(t, \tau)$ dans l'intervalle (t_0, t) , le long de la diagonale principale $t = \tau$ et que $F(t_0, t)$ est une fonction non décroissante de t , puisque chaque terme de la somme (2.2.5) est non négatif, car,

$$(2.2.6) \quad \Gamma(t_{N_k}, t_{N_k}) - \Gamma(t_{N_{k-1}}, t_{N_k}) - \Gamma(t_{N_k}, t_{N_{k-1}}) + \Gamma(t_{N_{k-1}}, t_{N_{k-1}}) = \|x(t_{N_k}) - x(t_{N_{k-1}})\|^2 \geq 0.$$

D'autre part, la continuité de $\mathbf{x}'(t/t)$ entraîne

$$(2.2.7) \quad \|\mathbf{x}'(t/t_0)\| \leq A < \infty$$

pour $t \geq t_0$ dans tout intervalle fini, d'après

$$\|\mathbf{x}'(t/t_0)\| \leq \|\mathbf{x}'(t/t)\|.$$

Le terme $x(t_{N_k}) - \mathbf{x}(t_{N_k}/t_{N_{k-1}})$ de la somme (2.2.3) peut s'écrire comme suit :

$$(2.2.8) \quad x(t_{N_k}) - \mathbf{x}(t_{N_k}/t_{N_{k-1}}) = [x(t_{N_k}) - x(t_{N_{k-1}})] - [\mathbf{x}(t_{N_k}/t_{N_{k-1}}) - x(t_{N_{k-1}})]$$

Et grâce à la relation d'orthogonalité

$$(2.2.9) \quad x(t_{N_k}) - \mathbf{x}(t_{N_k}/t_{N_{k-1}}) \perp \mathbf{x}(t_{N_k}/t_{N_{k-1}}) - x(t_{N_{k-1}})$$

le carré de sa norme est donné par :

$$(2.2.10) \quad \|x(t_{N_k}) - \mathbf{x}(t_{N_k}/t_{N_{k-1}})\|^2 = \|x(t_{N_k}) - x(t_{N_{k-1}})\|^2 - \|\mathbf{x}(t_{N_k}/t_{N_{k-1}}) - x(t_{N_{k-1}})\|^2$$

La somme $S_N(t_0, t)$ peut alors s'écrire comme la différence de deux sommes

$$S_N(t_0, t) = S'_N(t_0, t) - S''_N(t_0, t), \text{ où}$$

$$(2.2.11) \quad S'_N(t_0, t) = \sum_{k=1}^N [x(t_{N_k}) - x(t_{N_{k-1}})],$$

$$(2.2.12) \quad S''_N(t_0, t) = \sum_{k=1}^N [\mathbf{x}(t_{N_k}/t_{N_{k-1}}) - x(t_{N_{k-1}})];$$

$S'_N(t_0, t)$ étant égale à $x(t) - x(t_0)$, nous avons

$$S_N(t_0, t) = x(t) - x(t_0) - S''_N(t_0, t).$$

D'autre part, puisque $x(t)$ est à projections dérivables $\mathbf{x}(t_{N_k}/t_{N_{k-1}}) - x(t_{N_{k-1}})$ peut s'écrire pour $t_{N_k} - t_{N_{k-1}}$ suffisamment petit

$$\mathbf{x}(t_{N_k}/t_{N_{k-1}}) - x(t_{N_{k-1}}) = (t_{N_k} - t_{N_{k-1}}) [\mathbf{x}'(t_{N_{k-1}}/t_{N_{k-1}}) + \theta_{N_{k-1}}]$$

où $\theta_{N_{k-1}}$ est une v. a. appartenant à $\mathcal{E}_{N_{k-1}}(x)$ et tendant vers zéro en m. q. La somme S''_N peut alors s'écrire

$$(2.2.13) \quad S''_N(t_0, t) = \sum_{k=1}^N \mathbf{x}'(t_{N_{k-1}}/t_{N_{k-1}}) (t_{N_k} - t_{N_{k-1}}) + \sum_{k=1}^N (t_{N_k} - t_{N_{k-1}}) \theta_{N_{k-1}}.$$

La première somme tend en m. q., vers l'intégrale

$$\int_{t_0}^t \mathbf{x}'(s/s) ds$$

grâce à la continuité en m. q. de $\mathbf{x}'(t/t)$. Tandis que la deuxième somme de (2.2.13) tend, en m. q., vers zéro. En effet,

$$\left\| \sum_{k=1}^N (t_{N_k} - t_{N_{k-1}}) \theta_{N_{k-1}} \right\|^2 \leq (t - t_0)^2 \max \|\theta_{N_k}\|^2$$

et $\max \|\theta_{N_k}\|^2 \rightarrow 0$, lorsque $N \rightarrow \infty$, quel que soit le point considéré. La somme $S_N(t_0, t)$ a donc une limite en m. q. et elle est égale à

$$Z(t_0, t) = x(t) - x(t_0) - \int_{t_0}^t \mathbf{x}'(s/s) ds.$$

Grâce à (2.2.10), $\|S_N(t_0, t)\|^2$ peut s'écrire

$$\|S_N(t, t)\|^2 = \sum_{k=1}^N \|\mathbf{x}(t_{N_k}) - \mathbf{x}(t_{N_k}/t_{N_{k-1}})\|^2 = \sum_{k=1}^N \|x(t_{N_k}) - x(t_{N_{k-1}})\|^2 - \sum_{k=1}^N \|\mathbf{x}(t_{N_k}/t_{N_{k-1}}) - x(t_{N_{k-1}})\|^2$$

et comme, pour N assez grand,

$$\sum_{k=1}^N \|x(t_{N_k}/t_{N_{k-1}}) - x(t_{N_{k-1}})\|^2 < \sum_{w=1}^N (t_{N_k} - t_{N_{k-1}})^2 A^2 < (t - t_0) A^2 \delta^2$$

où $\delta = \max (t_{N_k} - t_{N_{k-1}})$ tend vers zéro lorsque $N \rightarrow \infty$ puisque alors $\delta \rightarrow 0$, nous avons

$$(2.2.14) \quad \lim_{N \rightarrow \infty} \sum_{k=0}^N \|x(t_{N_k}) - x(t_{N_{k-1}}/t_{N_{k-1}})\|^2 = \lim_{N \rightarrow \infty} \sum_{k=0}^N \|x(t_{N_k}) - x(t_{N_{k-1}})\|^2 = F(t_0, t) \quad \text{C.Q.F.D.}$$

Remarquons que nous définissons par $F(t_0, t)$ qui est non négative et non décroissante, une fonction additive d'intervalle. Nous pouvons poser

$$(2.2.15) \quad F(t_0, t) = F_1(t) - F_1(t_0),$$

en prenant $F_1(0) = 0$, on a évidemment

$$(2.2.16) \quad F_1(t) = \|Z_1(t)\|^2$$

1.2.3. Cas où la variance $F_1(t)$ de la fonction d'innovation $Z_1(t)$ est absolument continue.

Dans le cas où $F_1(t)$ est absolument continue par rapport à t , il existe une fonction $K_1(t)$ à carré du module sommable sur tout intervalle fini et définie pour presque tout t , telle que

$$(2.3.1) \quad F_1(t) - F_1(t_0) = \int_{t_0}^t |K_1(s)|^2 ds.$$

Ceci entraîne l'existence d'une fonction aléatoire $\xi_1(t)$ à accroissements homogènes non corrélés, de sorte que l'on a :

$$(2.3.2) \quad Z_1(t) - Z_1(t_0) = \int_{t_0}^t K_1(s) d\xi_1(s).$$

$K_1(t)$ étant défini par (2.3.1) à un facteur de module 1 près, nous conviendrons de choisir $K_1(t)$ non négatif. Dès lors, $\xi_1(t)$ est uniquement définie par

$$(2.3.3) \quad \xi_1(t) - \xi_1(t_0) = \int_{t_0}^t \frac{dZ_1(s)}{K_1(s)} \quad \text{si } K_1(t) \neq 0 \text{ presque partout}$$

et, dans ce cas, $\mathcal{L}_1(d\xi_1) = \mathcal{L}_1(dZ_1)$,

THÉOREME 2.3.1. — Dans les conditions du théorème 2.2.1, si la dérivée $\frac{\partial \Gamma(t, \tau)}{\partial t}$ existe et est continue pour $t < \tau$ et si elle a une limite continue en t lorsque $\tau \rightarrow t$, alors :

1° $\|Z_1(t)\|^2 = F_1(t)$ est absolument continue, 2° $K_1^2(t)$ est donné par

$$(2.3.4) \quad \begin{aligned} K_1^2(t) &= \lim_{h \uparrow 0} \frac{1}{h} \|x(t+h) - x(t)\|^2 = \lim_{h \downarrow 0} \frac{1}{h} \|x(t+h) - x(t+h/t)\|^2 \\ &= \lim_{h \downarrow 0} \frac{1}{h} \|x(t) - x(t-h)\|^2 = \lim_{h \downarrow 0} \frac{1}{h} \|x(t) - x(t/t-h)\|^2 \\ &= \lim_{h \downarrow 0} \frac{1}{2h} \|x(t+h) - x(t-h)\|^2 = \lim_{h \downarrow 0} \frac{1}{2h} \|x(t+h) - x(t+h/t-h)\|^2 \\ &= \lim_{\tau \downarrow t} \left[\frac{\partial \Gamma(t, \tau)}{\partial t} - \frac{\partial \bar{\Gamma}(t, \tau)}{\partial t} \right] \end{aligned}$$

Démonstration. — (Dans ce qui suit, $\frac{\partial^{m+n} \Gamma(u, v)}{\partial t^m \partial \tau^n}$ signifiera $\left[\frac{\partial^{m+n} \Gamma(t, \tau)}{\partial t^m \partial \tau^n} \right]_{\tau=v}^{t=u}$). Remarquons

tout d'abord que $\frac{\partial \Gamma(t, \tau)}{\partial \tau}$ existe et est continue pour $t < \tau$ et que sa limite est une fonction continue en t lorsque $\tau \rightarrow t$, d'après les conditions du théorème 2.2.1 puisque

$$\begin{aligned} \frac{\partial \Gamma(t, \tau)}{\partial t} &= \lim_{h \downarrow 0} \frac{1}{h} \langle x(t), x(\tau+h) - x(\tau) \rangle \\ &= \lim_{h \downarrow 0} \frac{1}{h} \langle x(t), x(\tau+h/t) - x(\tau/t) \rangle = \langle x(t), x'(\tau/t) \rangle. \end{aligned}$$

Les conditions du théorème précédent n'étant pas modifiées $Z_1(t)$ existe et la relation (2.2.4) est toujours valable. Nous allons montrer maintenant que la limite (2.2.5) est absolument continue.

En effet, en tenant compte des hypothèses faites, nous pouvons écrire

$$\Gamma(t+h, t+h) - \Gamma(t+h, t) = h \left[\frac{\partial \Gamma(t+h, t)}{\partial \tau} + \theta'_t(h) \right] \quad \text{et}$$

$$\Gamma(t, t+h) - \Gamma(t, t) = h \left[\frac{\partial \Gamma(t, t+0)}{\partial \tau} + \theta''_t(h) \right],$$

où θ' et θ'' tendent vers zéro avec h . Nous voyons que

$$\begin{aligned} (2.3.5) \quad & \lim_{h \downarrow 0} \frac{1}{h} \|x(t+h) - x(t)\|^2 \\ &= \lim_{h \downarrow 0} \frac{1}{h} \left\{ \|x(t+h) - x(t+h/t)\|^2 + \|x(t+h/t) - x(t)\|^2 \right\} \\ &= \lim_{h \downarrow 0} \frac{1}{h} \|x(t+h) - x(t+h/t)\|^2 = \frac{\partial \Gamma(t+0, t)}{\partial \tau} - \frac{\partial \Gamma(t, t+0)}{\partial \tau} \end{aligned}$$

La somme (2.2.5) peut alors s'écrire

$$\begin{aligned} (2.3.6) \quad & \sum_{k=1}^N \left\{ \Gamma(t_{N_k}, t_{N_k}) - \Gamma(t_{N_k}, t_{N_{k-1}}) - \Gamma(t_{N_{k-1}}, t_{N_k}) + \Gamma(t_{N_{k-1}}, t_{N_{k-1}}) \right\} \\ &= \sum_{k=1}^N (t_{N_k} - t_{N_{k-1}}) \left[\frac{\partial \Gamma(t_{N_{k-1}} + 0, t_{N_{k-1}})}{\partial \tau} - \frac{\partial \Gamma(t_{N_{k-1}}, t_{N_{k-1}} + 0)}{\partial \tau} \right] + \sum_{k=1}^N (t_{N_k} - t_{N_{k-1}}) \theta_{N_{k-1}} \end{aligned}$$

où θ_{N_k} tend vers zéro avec $(t_{N_{k+1}} - t_{N_k})$. Le second terme du dernier membre de (2.3.6) tend vers zéro, lorsque $N \rightarrow \infty$, puisque son module est majoré par $(t - t_0) \max |\theta_{N_k}|$ et que $\max |\theta_{N_k}| \rightarrow 0$ lorsque $N \rightarrow \infty$. La somme (2.3.6) tend visiblement vers

$$(2.3.7) \quad \int_{t_0}^t \left[\frac{\partial \Gamma(s+0, s)}{\partial \tau} - \frac{\partial \Gamma(s, s+0)}{\partial \tau} \right] ds = F_1(t) - F_1(t_0).$$

Par conséquent, $F_1(t)$ est absolument continue, avec une dérivée continue.

Par les relations (2.3.5) nous avons déjà démontré les deux premières relations (2.3.4). En remarquant que les hypothèses de continuité faites entraînent :

$$\begin{aligned} (2.3.8) \quad K_1^2(t) &= \frac{\partial \Gamma(t+0, t)}{\partial \tau} - \frac{\partial \Gamma(t, t+0)}{\partial \tau} = \frac{\partial \Gamma(t, t-0)}{\partial \tau} - \frac{\partial \Gamma(t-0, t)}{\partial \tau} \\ &= \frac{\partial \Gamma(t+0, t-0)}{\partial \tau} - \frac{\partial \Gamma(t-0, t+0)}{\partial \tau}, \end{aligned}$$

on peut démontrer, à partir d'un développement semblable à (2.3.6), les autres relations de (2.3.4) ; la dernière égalité de (2.3.4) étant une conséquence de la symétrie hermitienne de $\Gamma(t, \tau)$.

C.Q.F.D.

Grâce au dernier théorème que nous venons de démontrer, nous voyons que $K_1^2(t)$ peut être interprété comme une densité que nous appellerons *densité d'innovation* du processus $x(t)$.

La projection $x_1(t)$ de $x(t)$ sur $\mathcal{E}_t(d\xi_1)$ peut être représentée par l'intégrale

$$(2.3.9) \quad x_1(t) = \int_{-\infty}^t c_1(t, s) d\xi_1(s), \quad \text{avec}$$

$$(2.3.10) \quad \|x_1(t)\|^2 = \int_{-\infty}^t |c_1(t, s)|^2 ds,$$

où le noyau $c_1(t, s)$ peut être défini par

$$(2.3.11) \quad c_1(t, s) = \lim_{h \downarrow 0} \frac{1}{h} \langle x(t), \xi_1(s+h) - \xi_1(s) \rangle,$$

pour tout t et presque tout s .

Nous définissons par la représentation (2.3.9), une transformation linéaire entre $\mathcal{E}_t(x_1)$ et l'espace $L_2(-\infty, t)$ des fonctions à carré du module sommable dans $(-\infty, t]$. La transformation est isométrique et l'image de $\mathcal{E}_t(x_1)$ dans $L_2(-\infty, t)$ est un sous-espace de ce dernier. De ce fait, comme $\lim_{t \rightarrow \infty} L_1(-\infty, t) = \{0\}$, nous déduisons que $\mathcal{E}_{-\infty}(x_1) = \{0\}$. La représentation (2.3.9) est une représen-

tation canonique de $x_1(t)$ et en nous référant à la définition donnée au § 1.1 nous voyons que $x_1(t)$ est une composante non corrélée avec le passé infiniment lointain de $x(t)$. Comme

$$x(t) = x_1(t) + y(t),$$

où $y(t) \perp x_1(t)$ quels que soient t et τ , les deux composantes sont, toutes les deux, à projections dérivables comme $x(t)$. Ceci entraîne que $c_1(t, s)$ est dérivable en norme en tant que fonction de t .

1.2.4. Processus à projections analytiques.

DÉFINITIONS 2.4.1. — 1° *Analyticité* : Considérons un processus $x(t)$ indéfiniment dérivable en m. q. (par conséquent p. s.) ; désignons par $x^{(n)}(t)$ la dérivée n -ième. Nous dirons que $x(t)$ est un processus analytique si la somme

$$(2.4.1) \quad \sum_{n=0}^N \frac{(t-s)^n}{n!} x^{(n)}(s)$$

tend en m. q. vers $x(t)$, lorsque $N \rightarrow \infty$, quels que soient t et s .

Toutes les fois que nous poserons

$$(2.4.2) \quad x(t) = \sum_{n=0}^{\infty} \frac{(t-s)^n}{n!} x^{(n)}(s)$$

il sera sous-entendu que la somme du second membre existe comme limite en m. q. de la somme (2.4.1) et que cette limite est égale p. s. à $x(t)$.

D'après un théorème dû à M. Loève, pour qu'un processus $x(t)$ soit analytique, il faut et il suffit que sa fonction de covariance $\Gamma(t, \tau)$ soit analytique, dans tout voisinage de la diagonale principale $t = \tau$. Par conséquent, $\Gamma(t, \tau)$ peut être représentée, en fonction de ses dérivées en un point $t = \tau = s$ par le développement suivant

$$(2.4.3) \quad \begin{aligned} \Gamma(t, \tau) = \langle x(t), x(\tau) \rangle &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(t-s)^m}{m!} \cdot \frac{(\tau-s)^n}{n!} \left[\frac{\partial^{m+n} \Gamma(t, \tau)}{\partial t^m \partial \tau^n} \right]_{\tau=s} \Big|_{t=s} \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(t-s)^m}{m!} \cdot \frac{(\tau-s)^n}{n!} \langle x^{(m)}(s), x^{(n)}(s) \rangle. \end{aligned}$$

2° *Projections analytiques* : Nous dirons qu'un processus $x(t)$ est à projections analytiques si $\mathbf{x}(t/t_0)$ est analytique en t quels que soient t et t_0 ($t_0 < t$).

Les dérivées de tous les ordres par rapport à t , de $\mathbf{x}(t/t_0)$ existent en m. q. et p. s. et nous avons le développement infini suivant :

$$(2.4.4) \quad \mathbf{x}(t + h/t_0) = \sum_{n=0}^{\infty} \frac{h^n}{n!} \mathbf{x}^{(n)}(t/t_0)$$

pour tout $h > t_0 - t$ dans le sens de la définition (2.4.1).

LEMME 2.4.1. — Si $x(t)$ est à projections analytiques, toute projection de $\mathbf{x}(t/t_0)$ sur un sous-espace quelconque de $\mathcal{E}_{t_0}$ est encore analytique.

Soit P l'opérateur de projection sur un sous-espace quelconque de $\mathcal{E}_{t_0}$. Nous avons :

$$\| P \mathbf{x}(t + h/t_0) - \sum_{n=0}^N \frac{h^n}{n!} P \mathbf{x}^{(n)}(t/t_0) \|^2 \leq \| \mathbf{x}(t + h/t_0) - \sum_{n=0}^N \frac{h^n}{n!} \mathbf{x}^{(n)}(t/t_0) \|^2 = \| \sum_{n=N+1}^{\infty} \frac{h^n}{n!} \mathbf{x}^{(n)}(t/t_0) \|^2$$

Par conséquent
$$\lim_{N \rightarrow \infty} P \| \mathbf{x}(t + h/t_0) - \sum_{n=0}^N \frac{h^n}{n!} P \mathbf{x}^{(n)}(t/t_0) \|^2 = 0.$$

Désignons par $\mathbf{x}_+(t_0/t_0)$, la limite en m. q. de $\mathbf{x}(t_0 + h/t_0)$ et par $\mathbf{x}_-(t_0/t_0)$ la limite en m. q. de $\mathbf{x}(t_0/t_0 - h)$ quand $h \searrow 0$. On peut facilement montrer que ces deux limites existent et sont égales à $\mathbf{x}(t_0)$, lorsque $x(t)$ est continu en m. q. au point $t = t_0$. Cette définition peut être étendue aux dérivées de la projection.

Nous posons $\mathbf{x}'_+(t_0/t_0) = \lim_{h \downarrow 0} \text{m. q.} \frac{1}{h} [\mathbf{x}(t_0 + h/t_0) - \mathbf{x}(t_0)]$ pour la dérivée à droite, et $\mathbf{x}'_-(t_0/t_0) = \lim_{h \downarrow 0} \text{m. q.} \frac{1}{h} [\mathbf{x}(t_0/t_0 - h) - \mathbf{x}(t_0 - h)]$ pour la dérivée à gauche.

Nous supposons que, ces deux limites existent et sont égales et nous désignons par $\mathbf{x}''(t_0/t_0)$ la limite commune. En supposant que $\mathbf{x}'(t_0/t_0)$ est continu en m. q., nous pouvons définir $\mathbf{x}''(t_0/t_0)$ et remonter ainsi jusqu'à la dérivée n-ième $\mathbf{x}^{(n)}(t_0/t_0)$.

D'une façon générale, supposons que $\mathbf{x}^{(n-1)}(t_0/t_0)$ existe et qu'elle soit continue en m. q. Nous définissons $\mathbf{x}^{(n)}(t_0/t_0)$ par

$$(2.4.5) \quad \begin{aligned} \mathbf{x}^{(n)}(t_0/t_0) &= \lim_{h \downarrow 0} \text{m. q.} \frac{1}{h} [\mathbf{x}^{(n-1)}(t_0 + h/t_0) - \mathbf{x}^{(n-1)}(t_0/t_0)] \\ &= \lim_{h \downarrow 0} \text{m. q.} \frac{1}{h} [\mathbf{x}^{(n-1)}(t_0/t_0 - h) - \mathbf{x}^{(n-1)}(t_0 - h/t_0 - h)] \end{aligned}$$

si elle existe. Remarquons que si $\mathbf{x}^{(n)}(t_0/t_0)$ est continue en m. q. elle est égale à :

$$(2.4.6) \quad \mathbf{x}^{(n)}(t_0/t_0) = \lim_{h \downarrow 0} \text{m. q.} \mathbf{x}^{(n)}(t_0 + h/t_0) = \lim_{h \downarrow 0} \text{m. q.} \mathbf{x}^{(n)}(t_0/t_0 - h).$$

LEMME 2.4.2. — Pour que $x(t)$ soit à projections analytiques, il suffit que, les dérivées $\mathbf{x}^{(n)}(t_0/t_0)$ étant définies comme ci-dessus et toutes continues en m. q., l'on ait

$$(2.4.7) \quad \mathbf{x}(t/t_0) = \sum_{n=0}^{\infty} \frac{(t - t_0)^n}{n!} \mathbf{x}^{(n)}(t_0/t_0), \text{ quels que soient } t \text{ et } t_0 (t_0 < t).$$

La démonstration du lemme s'effectue en remarquant, comme pour le lemme précédent, que la projection conserve l'analyticité et que la projection de la dérivée $\mathbf{x}^{(n)}(t_0/t_0)$ sur un sous-espace quelconque de $\mathcal{E}_{t_0}$, est encore la dérivée n-ième de la projection de $x(t)$ sur ce sous-espace.

Ajoutons que, dans les conditions du lemme, tous les processus $\mathbf{x}^{(n)}(t/t)$ sont aussi à projections analytiques. Nous supposerons, dans ce qui suit, que tout processus à projections analytiques satisfait aux conditions du lemme 2.4.2.

THÉORÈME 2.4.1. — Soit $x(t)$, un processus aléatoire satisfaisant à l'hypothèse suivante :
Hypothèse H. — $\langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle$ est analytique en t, τ et s pour $s \leq t, s \leq \tau$. Alors :

— 1° $x(t)$ est la somme de deux processus aléatoires

$$(2.4.8) \quad x(t) = u(t) + v(t),$$

orthogonaux l'un à l'autre, en ce sens que $\langle u(t), v(\tau) \rangle = 0$ quels que soient t et τ ;

— 2° $u(t)$ est un processus purement non-déterministe qui admet la représentation canonique propre

$$(2.4.9) \quad u(t) = \int_{-\infty}^t c(t, s) d\xi(s)$$

où $\xi(t)$ est un processus à accroissements homogènes non corrélés ;

— 3° Si n est l'ordre de dérivabilité, en m. q. le plus élevé de $x(t)$, $c(t, s)$ est donné par

$$(2.4.10) \quad c(t, s) = \frac{1}{K_{n+1}(s)} \left[\frac{\partial^{n+1} \Gamma(t, s)}{\partial s^n} - \langle x(t), \mathbf{x}^{(n+1)}(s/s) \rangle \right] \quad \text{où}$$

$$(2.4.11) \quad K_{n+1}^2(t) = \lim_{\tau \searrow t} \left[\frac{\partial^{2n+1} \Gamma(t, \tau)}{\partial t^{n+1} \partial \tau^n} - \frac{\partial^{2n+1} \Gamma(t, \tau)}{\partial t^n \partial \tau^{n+1}} \right];$$

— 4° $v(t)$ est un processus analytique.

Remarque. — Le fait que $\langle \mathbf{x}(t/t), \mathbf{x}(\tau/t) \rangle = \Gamma(t, \tau)$ soit analytique pour $t \leq \tau$ n'entraîne nullement l'analyticité de $x(t)$, comme nous allons le voir. Compte tenu de la symétrie hermitienne de $\Gamma(t, \tau)$, on peut dire que $\Gamma(t, \tau)$ coïncide avec une surface analytique pour $t \leq \tau$ et une autre surface analytique pour $t \geq \tau$, c'est-à-dire qu'elle représente une surface analytique ayant une arête le long de la diagonale principale $t = \tau$, qui, elle-même, est une courbe analytique.

Démonstration. — La démonstration, qui donne aussi une méthode de construction de la représentation canonique (2.4.9), s'appuie sur les 8 propositions qui suivent.

PROPOSITION 1. — Tous les processus $\mathbf{x}^{(n)}(t/t)$ sont définis et sont à projections analytiques.

a) L'analyticité du produit scalaire $\langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle$, pour $s \leq t, \tau$, qui peut se mettre sous la forme :

$$(2.4.12) \quad \langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(t-s)^m}{m!} \cdot \frac{(\tau-s)^n}{n!} F_{m,n}(s),$$

où $F_{m,n}(s) \geq 0$, $F_{m,n} = \bar{F}_{n,m}$,

entraîne que les coefficients $F_{m,n}(s)$ sont tous fonctions analytiques de s . Les relations

$$\lim_{h,k \searrow 0} \langle \mathbf{x}(t+h/t), \mathbf{x}(t+k/t) \rangle = \lim_{h,k \searrow 0} \langle \mathbf{x}(t/t-h), \mathbf{x}(t/t-k) \rangle = F_{0,0}(t)$$

entraînent que $\mathbf{x}(t+0/t)$ et $\mathbf{x}(t/t-0)$ sont bien définis et égaux p. s.

Nous désignerons la valeur commune par $\mathbf{x}(t/t) = x(t)$. La fonction de covariance de $x(t)$ est alors donnée par

$$(2.4.13) \quad \begin{aligned} \Gamma(t, \tau) &= \langle \mathbf{x}(t/t), \mathbf{x}(\tau/\tau) \rangle = \langle \mathbf{x}(t/t), \mathbf{x}(\tau/t) \rangle \\ &= \sum_{n=0}^{\infty} \frac{(\tau-t)^n}{n!} F_{0,n}(t), \text{ pour } \tau \geq t. \end{aligned}$$

Nous allons, maintenant, montrer que $x(t)$ est continu en m. q. En effet :

$$\begin{aligned} \lim_{h \downarrow 0} \|x(t+h) - x(t)\|^2 &= \Gamma(t+h, t+h) - \Gamma(t+h, t) - \Gamma(t, t+h) + \Gamma(t, t) \\ &= \lim_{h \downarrow 0} \left\{ F_{0,0}(t+h) + F_{0,0}(t) - \sum_{n=0}^{\infty} \frac{h^n}{n!} [F_{0,n}(t) + F_{n,0}(t)] \right\} = 0. \end{aligned}$$

La continuité à gauche se démontre de la même manière.

Remarquer que l'analyticité des coefficients $F_{0,n}(t)$ n'entraîne nullement la dérivabilité en m. q. de $x(t)$, puisque

$$\|x(t+h) - x(t)\|^2 = \{F'_{0,0}(t) - [F_{0,1}(t) + F_{1,0}(t)]\} h + O(h^2)$$

$$\text{et que la limite } \lim_{h \downarrow 0} \frac{1}{h^2} \|x(t+h) - x(t)\|^2 \text{ peut ne pas exister.}$$

b) On peut définir de la même manière que $\mathbf{x}(t/t)$ tous les processus $\mathbf{x}^{(n)}(t/t)$. Puisque $\langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle$ est analytique, $\mathbf{x}^{(n)}(t/s)$ est analytique pour $t > s$ et nous avons, d'après (2.4.12)

$$(2.4.14) \quad \begin{aligned} \frac{\partial^{2n}}{\partial t^n \partial \tau^n} \langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle &= \langle \mathbf{x}^{(n)}(t/s), \mathbf{x}^{(n)}(\tau/s) \rangle \\ &= \sum_{p=n}^{\infty} \sum_{q=n}^{\infty} \frac{(t-s)^{p-n}}{(p-n)!} \cdot \frac{(\tau-s)^{q-n}}{(q-n)!} \cdot F_{p,q}(s). \end{aligned}$$

et grâce aux relations

$$\lim_{h,k \downarrow 0} \langle \mathbf{x}^{(n)}(t+h/t), \mathbf{x}^{(n)}(t+k/t) \rangle = \lim_{h,k \downarrow 0} \langle \mathbf{x}^{(n)}(t/t-h), \mathbf{x}^{(n)}(t/t-k) \rangle = F_{n,n}(t).$$

$\mathbf{x}^{(n)}(t+0/t)$ et $\mathbf{x}^{(n)}(t/t-0)$ sont bien définis et égaux p. s. La continuité de $\mathbf{x}^{(n)}(t/t)$ se démontre à partir du développement (2.4.14) comme celle de $x(t)$.

Considérons le développement

$$(2.4.15) \quad \langle \mathbf{x}^{(n)}(t/s), \mathbf{x}^{(n)}(t/s) \rangle = \sum_{p=m}^{\infty} \sum_{q=n}^{\infty} \frac{(t-s)^{p-m}}{(p-m)!} \cdot \frac{(\tau-s)^{q-n}}{(q-n)!} \cdot F_{p,q}(s)$$

qui est valable pour $t, \tau \geq s$. Nous voyons que

$$(2.4.16) \quad \langle \mathbf{x}^{(m)}(s/s), \mathbf{x}^{(n)}(s/s) \rangle = F_{m,n}(s) = \bar{F}_{n,m}(s).$$

Par conséquent, nous avons

$$(2.4.17) \quad \langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(t-s)^m}{m!} \cdot \frac{(\tau-s)^n}{n!} \langle \mathbf{x}^{(n)}(s/s), \mathbf{x}^{(m)}(s/s) \rangle.$$

Ce qui démontre la proposition 1.

PROPOSITION 2. — Tous les processus $\mathbf{x}^{(n)}(t/t)$, possèdent des densités d'innovation analytiques.

Les produits scalaires $\langle \mathbf{x}^{(n)}(s/s), \mathbf{x}^{(m)}(\tau/s) \rangle$ sont analytiques en s et τ pour $s \leq \tau$, puisque

$$(2.4.18) \quad \frac{\partial^m}{\partial \tau^m} \langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle = \sum_{n=0}^{\infty} \frac{(t-s)^n}{n!} \langle \mathbf{x}^{(n)}(s/s), \mathbf{x}^{(m)}(\tau/s) \rangle$$

est analytique en t , τ et s ($s \leq t, \tau$). Par conséquent, les coefficients $\langle \mathbf{x}^{(n)}(s/s), \mathbf{x}^{(m)}(\tau/s) \rangle$ sont analytiques en s et τ , pour $s \leq \tau$. Nous voyons, alors, que la fonction de covariance de $\mathbf{x}^{(n)}(t/t)$

$$\langle \mathbf{x}^{(n)}(t/t), \mathbf{x}^{(n)}(\tau/t) \rangle = \langle \mathbf{x}^{(n)}(t/t), \mathbf{x}^{(n)}(\tau/\tau) \rangle, \quad \tau \geq t$$

vérifie les conditions du théorème 2.3.1. $\mathbf{x}^{(n)}(t/t)$ possède donc une densité d'innovation. Nous allons maintenant montrer que cette densité d'innovation est une fonction analytique.

Puisque $\langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle$ est analytique en s , sa dérivée par rapport à s est aussi analytique.

$$(2.4.19) \quad \left\{ \begin{aligned} \frac{\partial}{\partial s} \langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle &= - \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} \frac{(t-s)^{m-1}}{(m-1)!} \cdot \frac{(\tau-s)^n}{n!} \cdot \langle \mathbf{x}^{(m)}(s/s), \mathbf{x}^{(n)}(s/s) \rangle \\ &- \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \frac{(t-s)^m}{m!} \cdot \frac{(\tau-s)^{n-1}}{(n-1)!} \cdot \langle \mathbf{x}^{(m)}(s/s), \mathbf{x}^{(n)}(s/s) \rangle \\ &+ \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(t-s)^m}{m!} \cdot \frac{(\tau-s)^n}{n!} \cdot \frac{\partial}{\partial s} \langle \mathbf{x}^{(m)}(s/s), \mathbf{x}^{(n)}(s/s) \rangle \end{aligned} \right.$$

Nous allons calculer $\frac{\partial}{\partial s} \langle \mathbf{x}(s/s), \mathbf{x}(s/s) \rangle$ de la manière suivante :

$$\frac{\partial}{\partial s} \langle \mathbf{x}^{(m)}(s/s), \mathbf{x}^{(n)}(s/s) \rangle = \lim_{h \downarrow 0} \frac{1}{h} [\langle \mathbf{x}^{(m)}(s+h/s+h), \mathbf{x}^{(n)}(s+h/s+h) \rangle - \langle \mathbf{x}^{(m)}(s/s), \mathbf{x}^{(n)}(s/s) \rangle]$$

(2.4.20)

$$\begin{aligned} &= \lim_{h \downarrow 0} \frac{1}{h} [\langle \mathbf{x}^{(m)}(s+h/s+h) - \mathbf{x}^{(m)}(s+h/s), \mathbf{x}^{(n)}(s+h/s+h) - \mathbf{x}^{(n)}(s+h/s) \rangle \\ &+ \langle \mathbf{x}^{(m)}(s+h/s) - \mathbf{x}^{(m)}(s/s), \mathbf{x}^{(n)}(s+h/s) - \mathbf{x}^{(n)}(s/s) \rangle + \langle \mathbf{x}^{(m)}(s/s), \mathbf{x}^{(n)}(s+h/s) - \mathbf{x}^{(n)}(s/s) \rangle] \\ &= K_{m,n}(s) + \langle \mathbf{x}^{(m+1)}(s/s), \mathbf{x}^{(n)}(s/s) \rangle + \langle \mathbf{x}^{(m)}(s/s), \mathbf{x}^{(n+1)}(s/s) \rangle, \quad \text{où nous avons posé} \end{aligned}$$

$$(2.4.21) \quad K_{m,n}(s) = \lim_{h \downarrow 0} \frac{1}{h} \langle \mathbf{x}^{(m)}(s+h/s+h) - \mathbf{x}^{(m)}(s+h/s), \mathbf{x}^{(n)}(s+h/s) - \mathbf{x}^{(n)}(s+h/s) \rangle;$$

quantité qui peut être considérée comme un produit scalaire de densités d'innovations. En particulier $K_{n,n}(s)$ est effectivement la densité d'innovation de $\mathbf{x}^{(n)}(t/t)$. En reportant l'expression (2.4.20) de la dérivée, dans (2.4.19) nous trouvons le développement en série

$$(2.4.22) \quad \frac{\partial}{\partial s} \langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(t-s)^m}{m!} \cdot \frac{(\tau-s)^n}{n!} \cdot K_{m,n}(s)$$

que nous appellerons la *série des innovations*.

Puisque la somme de cette série est analytique en s ($s \leq t, \tau$), il en sera de même de tous les coefficients $K_{m,n}(s)$ et, en particulier des coefficients $K_{n,n}(s)$. Ce qui démontre la proposition 2.

Nous supposons, pour l'instant, que $K_{0,0}(t)$ n'est pas identiquement nul. Par conséquent, $K_1^2(t) = K_{0,0}(t)$ ne s'annulant qu'en des points isolés, à cause de son analyticit   et les conditions du th  or  me 2.3.1 se trouvant satisfaites, $\xi_1(t)$ existe et est uniquement d  fini.

PROPOSITION 3. — Le noyau $c_1(t, s)$ correspondant    la projection $x_1(t)$ de $x(t)$ sur $\mathcal{E}_t(d\xi_1)$, dans sa repr  sentation canonique

$$(2.4.23) \quad x_1(t) = \int_{-\infty}^t c_1(t, s) d\xi_1(s), \quad \text{est donn   par}$$

$$(2.4.24) \quad c_1(t, s) = \frac{1}{K_1(s)} \left[\frac{\partial}{\partial s} \Gamma(t, s) - \langle x(t), \mathbf{x}'(s/s) \rangle \right]$$

pour tout t et s ($s \leq t$).

Nous allons, tout d'abord, montrer que

$$\begin{aligned}
 (2.4.25) \quad & \lim_{h \downarrow 0} \frac{1}{h^2} \|x(t+h) - \mathbf{x}(t+h/t) - [Z_1(t+h) - Z_1(t)]\|^2 = \\
 & \lim_{h \downarrow 0} \frac{1}{h^2} \|x(t) - \mathbf{x}(t/t-h) - [Z_1(t) - Z_1(t-h)]\|^2 = \\
 & \lim_{h \downarrow 0} \frac{1}{4h^2} \|x(t+h) - \mathbf{x}(t+h/t-h) - [Z_1(t+h) - Z_1(t-h)]\|^2 = 0.
 \end{aligned}$$

Considérons seulement la première relation. Nous pouvons écrire en prenant une subdivision $\{t_{N_j}\}$ de l'intervalle $[t, t+h]$

$$\begin{aligned}
 (2.4.26) \quad & x(t+h) - \mathbf{x}(t+h/t) - [Z(t+h) - Z(t)] \\
 & = \lim_{N \rightarrow \infty} \text{m. q.} \sum_{j=1}^N \left\{ [\mathbf{x}(t+h/t_{N_j}) - \mathbf{x}(t+h/t_{N_{j-1}})] - [x(t_{N_j}) - \mathbf{x}(t_{N_j}/t_{N_{j-1}})] \right\} \\
 & = \lim_{N \rightarrow \infty} \text{m. q.} \sum_{j=1}^N [\mathbf{x}(t+h/t_{N_j}) - x(t_{N_j})] - [\mathbf{x}(t+h/t_{N_{j-1}}) - \mathbf{x}(t_{N_j}/t_{N_{j-1}})].
 \end{aligned}$$

Nous avons grâce à l'analyticité des projections

$$(2.4.27) \quad \mathbf{x}(t+h/t_{N_j}) - x(t_{N_j}) = \sum_{n=0}^{\infty} \frac{(t+h-t_{N_j})^n}{n!} \mathbf{x}^{(n)}(t_{N_j}/t_{N_j}).$$

Comme $\mathbf{x}(t+h/t_{N_{j-1}}) - \mathbf{x}(t_{N_j}/t_{N_{j-1}})$ est la projection de (2.4.27) sur $\mathcal{E}_{t_{N_{j-1}}}$ nous pouvons écrire

$$(2.4.28) \quad \mathbf{x}(t+h/t_{N_{j-1}}) - \mathbf{x}(t_{N_j}/t_{N_{j-1}}) = \sum_{n=1}^{\infty} \frac{(t+h-t_{N_j})^n}{n!} \mathbf{x}^{(n)}(t_{N_j}/t_{N_{j-1}}).$$

Et en prenant la différence de (2.4.27) et de (2.4.28), il vient :

$$\begin{aligned}
 & [\mathbf{x}(t+h/t_{N_j}) - x(t_{N_j})] - [\mathbf{x}(t+h/t_{N_{j-1}}) - \mathbf{x}(t_{N_j}/t_{N_{j-1}})] = \\
 & = \sum_{n=1}^{\infty} \frac{(t+h-t_{N_j})^n}{n!} [\mathbf{x}^{(n)}(t_{N_j}/t_{N_j}) - \mathbf{x}^{(n)}(t_{N_j}/t_{N_{j-1}})]
 \end{aligned}$$

En reportant cette dernière expression dans (2.4.26)

$$\begin{aligned}
 & x(t+h) - \mathbf{x}(t+h/t) - [Z(t+h) - Z(t)] \\
 & = \lim_{N \rightarrow \infty} \text{m. q.} \sum_{j=1}^N \left\{ \sum_{n=1}^{\infty} \frac{(t+h-t_{N_j})^n}{n!} [\mathbf{x}^{(n)}(t_{N_j}/t_{N_j}) - \mathbf{x}^{(n)}(t_{N_j}/t_{N_{j-1}})] \right\}
 \end{aligned}$$

Tenant compte de l'orthogonalité des termes de la série (2.4.26), nous avons

$$\begin{aligned}
 (2.4.29) \quad & \|x(t+h) - \mathbf{x}(t+h/t) - [Z(t+h) - Z(t)]\|^2 = \lim_{N \rightarrow \infty} \sum_{j=1}^N \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{(t+h-t_{N_j})^{m+n}}{m!n!} \cdot U_{t_{N_j}}^{m,n}, \\
 & \text{où } U_{t_{N_j}}^{m,n} = \langle \mathbf{x}^{(m)}(t_{N_j}/t_{N_j}) - \mathbf{x}^{(m)}(t_{N_j}/t_{N_{j-1}}), \mathbf{x}^{(n)}(t_{N_j}/t_{N_{j-1}}) - \mathbf{x}^{(n)}(t_{N_j}/t_{N_{j-1}}) \rangle
 \end{aligned}$$

Mais nous pouvons écrire, grâce à l'existence de la série des innovations, pour N assez grand,

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{(t+h-t_{N_j})^{m+n}}{m!n!} U_{t_{N_j}}^{m,n} = (t_{N_j} - t_{N_{j-1}}) \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{(t+h-t_{N_j})^{m+n}}{m!n!} \cdot [K_{m,n}(t_{N_j}) + \theta_{N_j}^{m,n}]$$

où $\theta_{N_j}^{m,n} \rightarrow 0$ avec $t_{N_j} - t_{N_{j-1}}$. En reportant cette série dans (2.4.29) il vient

$$\begin{aligned}
 & \|x(t+h) - \mathbf{x}(t+h/t) - [Z(t+h) - Z(t)]\|^2 \\
 & = \lim_{N \rightarrow \infty} \sum_{j=1}^N (t_{N_j} - t_{N_{j-1}}) \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{(t+h-t_{N_j})^{m+n}}{m!n!} [K_{m,n}(t_{N_j}) + \theta_{N_j}^{m,n}] \\
 & = \int_t^{t+h} \left[\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{(t+h-s)^{m+n}}{m!n!} K_{m,n}(s) \right] ds = \int_t^{t+h} (t+h-s)^2 \left[\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{(t+h-s)^{m+n-2}}{m!n!} K_{m,n}(s) \right] ds.
 \end{aligned}$$

La quantité entre crochets est bornée et elle est même analytique en s , d'après l'analyticité en s de la série des innovations (2.4.22). Par conséquent, il existe un nombre A fini, tel que

$$\|x(t+h) - \mathbf{x}(t+h/t) - [Z(t+h) - Z(t)]\|^2 \leq A \int_t^{t+h} (t+h-s)^2 ds = A h^3/3.$$

Ce qui démontre la première relation (2.4.25). La démonstration des deux autres relations peut s'effectuer de la même manière. Puisque $K_1(t) \neq 0$, presque partout, nous pouvons écrire, d'après (2.3.11).

$$\begin{aligned} c_1(t, s) &= \lim_{h \downarrow 0} \frac{1}{h} \langle x(t), \xi_1(s+h) - \xi_1(s) \rangle = \frac{1}{K_1(s)} \lim_{h \downarrow 0} \frac{1}{h} \langle x(t), Z_1(s+h) - Z_1(s) \rangle \\ &= \frac{1}{K_1(s)} \lim_{h \downarrow 0} \frac{1}{h} \langle x(t), x(s+h) - \mathbf{x}(s+h/s) \rangle = \frac{1}{K_1(s)} \lim_{h \downarrow 0} \left[\frac{1}{h} \langle x(t), x(s+h) - x(s) \rangle \right. \\ &\quad \left. - \frac{1}{h} \langle x(t), \mathbf{x}(s+h/s) - x(s) \rangle \right] = \frac{1}{K_1(s)} \left[\frac{\partial \Gamma(t, s)}{\partial s} - \langle x(t), \mathbf{x}'(s/s) \rangle \right]. \end{aligned}$$

PROPOSITION 4. — $x_1(t)$ donné par (2.4.23) vérifie aussi l'hypothèse H.

$x_1(t)$ est un processus à projections analytiques d'après le lemme 2.4.1. Ce fait entraîne l'analyticité de $c_1(t, s)$ comme fonction de t pour presque tout $s (s < t)$, ce qui est aussi une conséquence de l'expression (2.4.24) de $c_1(t, s)$. Le noyau correspondant à $\mathbf{x}_1^{(n)}(t/t)$ dans sa représentation intégrale est

$$\begin{aligned} \frac{\partial^n c_1(t, s)}{\partial t^n} \text{ et est donné par } \frac{\partial^n}{\partial t^n} c_1(t, s) &= \frac{1}{K_1(s)} \left[\frac{\partial^{n+1}}{\partial t^n \partial s} \Gamma(t, s) - \frac{\partial^n}{\partial t^n} \langle x(t), \mathbf{x}'(s/s) \rangle \right] \\ (2.4.30) \quad &= \frac{1}{K_1(s)} \left[\frac{\partial^{n+1}}{\partial t^n \partial s} \Gamma(t, s) - \langle \mathbf{x}^{(n)}(t/s), \mathbf{x}'(s/s) \rangle \right] \end{aligned}$$

$c_1(t, s) \cdot \bar{c}_1(\tau, s)$ est intégrable par rapport à s , puisque $\|x_1(s)\|^2 < \infty$. Ceci entraîne que

$$(2.4.31) \quad c_1(t, s) \bar{c}_1(\tau, s) = \frac{\left[\frac{\partial \Gamma(t, s)}{\partial s} - \langle x(t), \mathbf{x}'(s/s) \rangle \right] \left[\frac{\Gamma(s, \tau)}{\partial s} - \langle \mathbf{x}'(s/s), x(\tau) \rangle \right]}{K_1^2(t)}$$

est une fonction analytique de t, τ et de $s (s \leq t, \tau)$. Car, il en est évidemment ainsi si $K_1^2(s)$ ne s'annule jamais, puisque le numérateur et le dénominateur de (2.4.31) sont analytiques en s . Mais si $K_1^2(s)$ s'annule pour des valeurs isolées de s , les zéros étant au moins d'ordre deux, puisque $K_1^2(s)$ est non négatif ; ou bien le numérateur de (2.4.31) s'annule aux mêmes points que $K_1^2(s)$, l'ordre des zéros étant au moins le même que pour $K_1^2(s)$, auquel cas (2.4.31) est encore analytique ; ou bien le numérateur ne s'annule pas en même temps que $K_1^2(s)$, alors (2.4.31) ne serait plus intégrable en s et ceci est impossible.

Comme l'intégrale indéfinie de l'expression (2.4.31) par rapport à s existe et que (2.4.31) est analytique en t, τ et $s (s \leq t, \tau)$; alors, son intégrale

$$\langle \mathbf{x}_1(t/s), \mathbf{x}_1(\tau/s) \rangle = \int_{-\infty}^s c_1(t, u) \cdot \bar{c}_1(\tau, u) du$$

est aussi analytique en t, τ et $s (s \leq t, \tau)$.

PROPOSITION 5. — Le processus $y(t) = x(t) - x_1(t)$ vérifie l'hypothèse H, sa densité d'innovation est identiquement nulle et il est dérivable au moins une fois.

La continuité et l'analyticité du noyau $c_1(t, s)$ (2.4.24) entraînent que

$$\mathcal{E}_t(x_1) = \mathcal{E}_t(d\xi_1).$$

Par conséquent, la représentation (2.4.23) de $x_1(t)$ est une représentation canonique propre (cf. Lévy [3]). Nous avons, d'après la relation d'orthogonalité

$$y(t) \perp \mathcal{E}(x_1),$$

$\mathbf{x}(t/s) = \text{Projection de } x_1(t) \text{ sur } \mathcal{E}_s(x_1) + \text{Projection de } y(t) \text{ sur } \mathcal{E}_s(y).$

$= \text{Projection de } [x_1(t) + y(t)] \text{ sur } \mathcal{E}_t(x_1) \oplus \mathcal{E}_t(y) = \mathcal{E}_t(x). \quad *$

Comme alors $x_1(t)$ vérifie l'hypothèse H, $y(t)$ aussi la vérifie.

D'autre part, nous avons vu que $x_1(t)$ vérifie les conditions du théorème 2.3.1, sa densité d'innovation est donnée par

$$\lim_{h \downarrow 0} \frac{1}{h} \|x_1(t+h) - \mathbf{x}_1(t+h/t)\|^2 = \lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} |c_1(t+h, s)|^2 ds.$$

* $\oplus$ désigne la somme directe des espaces de Hilbert.

Puisque $c_1(t, s)^2$ est continu par rapport à t et à s ($s \leq t$), nous avons

$$\lim_{h \downarrow 0} \frac{1}{h} \|x_1(t+h) - x_1(t+h/t)\|^2 = |c_1(t, t)|^2.$$

Mais, d'après l'expression de $c_1(t, s)$, il vient

$$(2.4.32) \quad |c_1(t, t)|^2 = \frac{1}{K_1^2(t)} \left| \left[\frac{\partial \Gamma(t, s)}{\partial s} \right]_{s \nearrow t} - \langle x(t), \mathbf{x}'(t/t) \rangle \right|^2.$$

Comme nous avons

$$\frac{\partial \Gamma(t, \tau)}{\partial \tau} = \langle x(t), \mathbf{x}'(t/t) \rangle \quad \tau \geq t,$$

et que cette quantité est continue par rapport à τ , pour $\tau \geq t$,

$$\langle x(t), \mathbf{x}'(t/t) \rangle = \left[\frac{\partial \Gamma(t, \tau)}{\partial \tau} \right]_{\tau \searrow t} = \left[\frac{\partial \bar{\Gamma}(t, s)}{\partial t} \right]_{s \nearrow t}$$

Et en reportant cette dernière expression dans (2.4.32), nous trouvons

$$(2.4.33) \quad |c_1(t, t)|^2 = \frac{1}{K_1^2(t)} \left| \left[\frac{\partial \Gamma(t, s)}{\partial s} \right]_{s \nearrow t} - \left[\frac{\partial \bar{\Gamma}(t, s)}{\partial t} \right]_{s \nearrow t} \right|^2 = K^2(t).$$

Par conséquent, $x(t)$ et $x_1(t)$ ayant les mêmes densités d'innovation, la densité d'innovation de $y(t)$ est identiquement nulle.

Soient $\Gamma_1(t, \tau)$ la fonction de covariance de $x_1(t)$ et $\Gamma_y(t, \tau)$, celle de $y(t)$. Nous avons

$$\Gamma_y(t, \tau) = \Gamma(t, \tau) - \Gamma_1(t, \tau) \quad \text{et}$$

$$(2.4.34) \quad \lim_{\tau \downarrow t} \left[\frac{\partial \Gamma_y(t, \tau)}{\partial t} - \frac{\partial \bar{\Gamma}_y(t, \tau)}{\partial \tau} \right] = \lim_{h \downarrow 0} \frac{1}{h} \|y(t+h) - \mathbf{y}(t+h/t)\|^2 = 0.$$

Nous allons montrer maintenant que cette dernière relation entraîne la dérivabilité en m. q. de $y(t)$. Nous avons, d'une part,

$$(2.4.35) \quad \frac{\partial \Gamma_y(t, \tau)}{\partial \tau} = \langle y(t), \mathbf{y}'(\tau/t) \rangle, \quad \text{pour } \tau > t$$

et d'autre part, pour $\tau > t$,

$$\begin{aligned} \frac{\partial \Gamma_y(t, \tau)}{\partial t} &= \lim_{h \downarrow 0} \frac{1}{h} \langle y(t+h) - y(t), y(\tau) \rangle = \lim_{h \downarrow 0} \frac{1}{h} \langle y(t+h) - \mathbf{y}(t+h/t), y(\tau) \rangle \\ &\quad + \lim_{h \downarrow 0} \frac{1}{h} \langle \mathbf{y}(t+h/t) - y(t), y(\tau) \rangle \\ &= \lim_{h \downarrow 0} \frac{1}{h} \langle y(t+h) - \mathbf{y}(t+h/t), y(\tau) \rangle + \langle \mathbf{y}'(t/t), y(\tau) \rangle. \end{aligned}$$

La dernière limite est nulle quels que soient t et τ , car

$$\begin{aligned} \lim_{h \downarrow 0} \left| \frac{1}{h} \langle y(t+h) - \mathbf{y}(t+h/t), y(\tau) \rangle \right|^2 &\leq \lim_{h \downarrow 0} \frac{1}{h} \|y(t+h) - \mathbf{y}(t+h/t)\|^2 \\ &\quad \frac{1}{h} \|\mathbf{y}(t/t) - \mathbf{y}(\tau/t)\|^2 = 0 \end{aligned}$$

puisque $\lim_{h \downarrow 0} \frac{1}{h} \|\mathbf{y}(\tau/t) - \mathbf{y}(t/t)\|^2 < \infty$, d'après (2.4.19) (il en est de même pour $x_1(t)$)

et que $\lim_{h \downarrow 0} \frac{1}{h} \|y(t+h) - \mathbf{y}(t+h/t)\|^2 = 0$ d'après (2.4.34). Finalement

$$(2.4.36) \quad \frac{\partial \Gamma_y(t, \tau)}{\partial t} = \langle \mathbf{y}'(t/t), y(\tau) \rangle.$$

Par conséquent, la comparaison de (2.4.35) et (2.4.36) montre que la dérivée partielle $\frac{\partial \Gamma_y(t, \tau)}{\partial t}$ est continue à la traversée de la diagonale principale $t = \tau$.

En dérivant encore une fois (2.4.35) et (2.4.26), nous trouvons

$$(2.4.37) \quad \frac{\partial^2 \Gamma_y(t, \tau)}{\partial \tau^2} = \langle y(t), y''(\tau/t) \rangle, \quad (\tau > t), \quad \text{et}$$

$$(2.4.38) \quad \begin{aligned} \frac{\partial^2 \Gamma_y(t, \tau)}{\partial t^2} &= \lim_{h \downarrow 0} \frac{1}{h} \langle y(t+h/t+h) - y(t/t), y(\tau) \rangle \\ &= \lim_{h \downarrow 0} \frac{1}{h} \langle y'(t+h/t+h) - y'(t+h/t), y(\tau) \rangle + \langle y''(t/t), y(\tau) \rangle \end{aligned}$$

La limite du second membre existe uniformément en τ , puisqu'il en est de même pour $\frac{\partial^2 \Gamma_y(t, \tau)}{\partial t^2}$ et $\langle y''(t/t), y(\tau) \rangle$ ($t < \tau$), et elle est continue par rapport à τ uniformément en t . Par conséquent nous sommes en droit d'écrire

$$\begin{aligned} &\lim_{h \downarrow 0} \left[\lim_{h \downarrow 0} \frac{1}{h} \langle y'(t+h/t+h) - y'(t+h/t), y(\tau) \rangle \right] \\ &= \lim_{h \downarrow 0} \frac{1}{h} \langle y'(t+h/t+h) - y'(t+h/t), y(t+h/t+h) - y(t+h/t) \rangle. \quad \text{Donc} \\ &\lim_{h \downarrow 0} \left| \frac{1}{h} \langle y'(t+h/t+h) - y'(t+h/t), y(t+h/t+h) - y(t+h/t) \rangle \right|^2 \\ (2.4.39) \quad &\leq \lim_{h \downarrow 0} \frac{1}{h} \|y'(t+h/t+h) - y'(t+h/t)\|^2 \cdot \frac{1}{h} \|y(t+h/t+h) - y(t+h/t)\|^2 = 0 \end{aligned}$$

puisque $\lim_{h \downarrow 0} \frac{1}{h} \|y'(t+h/t+h) - y'(t+h/t)\|^2$ existe, étant la densité d'innovation de $y'(t/t)$, et que

$$\lim_{h \downarrow 0} \frac{1}{h} \|y(t+h/t+h) - y(t+h/t)\|^2 = 0.$$

Nous voyons alors, compte tenu de (2.4.37), (2.4.38) et (2.4.39), que la dérivée seconde $\frac{\partial^2 \Gamma_y(t, \tau)}{\partial t^2}$ n'a pas de discontinuité à la traversée de la diagonale principale $t = \tau$.

On peut montrer, de la même manière, la continuité de $\frac{\partial^n \Gamma_y(t, \tau)}{\partial t^n}$ quel que soit n . Alors compte tenu de la continuité de ces dérivées, nous pouvons écrire

$$\begin{aligned} &\Gamma_y(t+h, t+k) - \Gamma_y(t+h, t) - \Gamma_y(t, t+k) + \Gamma_y(t, t) \\ &= \sum_{k=1}^{\infty} \frac{k^n}{n!} \left[\frac{\partial^n}{\partial \tau^n} \Gamma_y(t+h, t) - \frac{\partial^n}{\partial \tau^n} \Gamma_y(t, t) \right] \\ &= \sum_{k=1}^{\infty} \frac{hk^n}{n!} \left[\frac{\partial^{n+1} \Gamma_y(t+\theta_n h, \tau)}{\partial t \partial \tau^n} \right]_{\tau=t} \quad \text{où } 0 \leq \theta \leq 1. \text{ Par conséquent,} \end{aligned}$$

$\lim_{h, k \rightarrow 0} \frac{1}{h, k} [\Gamma_y(t+h, t+k) - \Gamma_y(t+h, t) - \Gamma_y(t, t+k) + \Gamma_y(t, t)] = \lim_{h \rightarrow 0} \left[\frac{\partial^2 \Gamma_y(t+h, \tau)}{\partial t \partial \tau} \right]_{\tau=t}$ ce qui prouve la dérivabilité en m. q. de $y(t)$. D'autre part, les relations (2.4.35) et (2.4.36) montrent que la dérivée en m. q. $y'(t)$ de $y(t)$ est égale p. s. à $y'(t/t)$.

PROPOSITION 6. — $x_1(t)$ est la seule composante non-déterministe de $x(t)$.

Supposons, au départ, que la densité d'innovation de la dérivée $y'(t)$ de $y(t)$ soit différente de zéro. $y'(t)$ vérifiant, à son tour, l'hypothèse H, sa densité d'innovation sera analytique. On peut donc, au même titre que pour $x(t)$, attacher à $y'(t)$ une fonction aléatoire $\xi_2(t)$ à accroissements homogènes non corrélés, définie exactement comme $\xi_1(t)$, à partir des innovations de $y'(t)$. La projection de $x(t)$ (ou de $y(t)$) sur $\mathcal{E}_t(d\xi_2)$, peut être représentée par

$$x_2(t) = \int_{-\infty}^t c_2(t, s) d\xi_2(s),$$

où $c_2(t, s)$ est nul, comme $c_1(t, s)$ pour $t < s$ et à carré du module sommable par rapport à s .

Nous allons montrer maintenant, qu'il est impossible que $y'(t)$ ait une densité d'innovation différente de zéro et qu'il existe une telle composante $x_2(t)$ de $x(t)$, en faisant appel aux propriétés des espaces de Hilbert à *noyau reproduisant* (cf. Aronszajn et Hida).

Soit $\Gamma(t, \tau)$ une fonction de covariance que nous supposons définie dans tout le plan (t, τ) . Il existe un espace de Hilbert $\mathcal{H}$ ayant les propriétés suivantes :

1° $\Gamma(t, \tau)$ appartient à $\mathcal{H}$,

2° $\langle f(t), \Gamma(t, \tau) \rangle_{\Gamma} = f(\tau)$ pour tout $f \in \mathcal{H}$ comme fonction de t ,

3° $\mathcal{H}$ est un espace linéaire fermé (dans la topologie définie par la norme $\|\cdot\|_{\Gamma}$) engendré par $\{\Gamma(\cdot, \tau); -\infty < \tau < \infty\}$.

$\Gamma(t, \tau)$ s'appelle le *noyau reproduisant* de cet espace de Hilbert.

Nous supposons que $\mathcal{E}_{-\infty}(x) = \{0\}$ et que $\mathcal{H}$ est séparable. Soit $\mathcal{H}_t$ le sous-espace de $\mathcal{H}$ engendré par $\{\Gamma(\cdot, s); s \leq t\}$ et désignons par $E(t)$ la résolution de l'identité :

$$\mathcal{H}_t^* = E(t) \mathcal{H}, \text{ où } \mathcal{H}_t^* = \bigcap_n \mathcal{H}_{t+1/n}$$

Alors, il existe dans $\mathcal{H}$ deux ensembles $\{f_n(t)\}$ ($n = 1, 2, \dots$) et $g^{k,j}$ ($k, j = 1, 2, \dots$) satisfaisant aux conditions suivantes :

1° Pour tout intervalle Δ_1, Δ_2

$$\langle \Delta_1 E(s) f_m(\cdot), \Delta_2 E(s) f_n(\cdot) \rangle = 0, \text{ pour } m \neq n,$$

$$\text{avec } \Delta E = E(b) - E(a) \text{ pour } \Delta = (a, b);$$

2° Si Δ_1 et Δ_2 sont deux intervalles disjoints

$$\langle \Delta_1 E(s) f_n(\cdot), \Delta_2 E(s) f_n(\cdot) \rangle = 0,$$

3° Pour tout n , $\mu_n(s) = \|E(s) f_n(\cdot)\|^2$ est continue, non décroissante et $\mu_{n+1} \leq \mu_n$ (considérées comme mesures).

4° $g^{k,j}$ est le vecteur propre de l'opérateur auto-adjoint

$$H = \int t dE(t)$$

correspondant à la valeur propre t_k ($j = 1, 2, \dots$).

$$5^\circ \quad \mathcal{H} = \mathfrak{M} \oplus \mathfrak{N}$$

où $\mathfrak{M}$ et $\mathfrak{N}$ sont définis de la façon suivante

$$\mathfrak{M} = \sum_n \oplus \mathfrak{M}(f_n), \quad \text{avec}$$

$$\mathfrak{M}(f_n) = \left\{ f; f(t) = \int \varphi(s) dE(s) f_n(t), \varphi \in L_2(d\mu_n) \right\};$$

$L_2(d\mu_n)$ étant l'espace de Hilbert des fonctions à carré du module sommable par rapport à la mesure μ_n .

$$\mathfrak{N} = \sum_{t_k} \sum_j \oplus \mathfrak{N}(g^{k,j})$$

où $\mathfrak{N}(g^{k,j})$ est le sous-espace à une dimension engendré par $g^{k,j}$.

A remarquer que la relation

$$(2.4.40) \quad f(t) = \int \varphi(s) dE(s) f_n(t)$$

est une transformation isométrique biunivoque entre $\mathfrak{M}(f_n)$ et $L_2(d\mu_n)$.

Grâce à ces propriétés, la fonction de covariance $\Gamma(t, \tau)$ admet la représentation suivante (cf. Hida)

$$(2.4.41) \quad \Gamma(t, \tau) = \sum_n \int_{-\infty}^t g_n(t, s) dE(s) f_n(\tau) + \sum_{t_k \leq t} b_k^1(t) \cdot \bar{b}_k^1(\tau), \quad \text{où}$$

$$(2.4.42) \quad g_n(t, s) = \frac{dE(s) \cdot f_n(t)}{d\mu_n(s)} = \frac{\langle \Gamma(\cdot, t), dE(s) f_n(\cdot) \rangle}{d\mu_n(s)}$$

pour presque tout s , μ_n — mesure.

Nous voyons que la représentation canonique généralisée (2.1.9) n'est qu'une conséquence de la représentation (2.4.41) de la covariance.

Supposons, maintenant, que $\mu_1(s)$ et $\mu_2(s)$ se réduisent à la mesure de Lebesgue, comme dans le cas où $x(t)$ aurait deux composantes $x_1(t)$ et $x_2(t)$ dont nous avons donné la représentation ci-dessus. Nous aurons :

$$\frac{dE(s) \cdot f_i(t)}{ds} = c_i(t, s) \quad (i = 1, 2) \quad \text{pour presque tout } s; \text{ et}$$

$$\mathfrak{N}(f_i) = \left\{ f; f(t) = \int_{-\infty}^t \varphi(s) c_i(t, s) ds, \varphi \in L_2(-\infty, \infty) \right\}, \text{ avec}$$

$$(2.4.43) \quad \mathfrak{N}(f_1) \perp \mathfrak{N}(f_2).$$

Nous allons montrer maintenant, que cette dernière relation n'est possible que si $c_2(t, s) = 0$ pour presque tout s . En effet, $c_1(t, s)$ et $c_2(t, s)$ étant toutes les deux à carré du module par rapport à s , l'élément

$$(2.4.44) \quad g(t) = \int_{-\infty}^t \bar{c}_1(t, s) \bar{c}_2(t, s) ds$$

devrait appartenir à la fois à $\mathfrak{N}(f_1)$ et à $\mathfrak{N}(f_2)$. Mais ceci n'est possible d'après (2.4.43) que si $\|g\|_{\Gamma}^2 = 0$. Par conséquent, d'après la relation

$$|g|^2 = |\langle g(\cdot), \Gamma(\cdot, \tau) \rangle|^2 \leq \|g\|_{\Gamma}^2 \cdot \|\Gamma(\cdot, \tau)\|_{\Gamma}^2$$

$g(t)$ doit être identiquement nulle. Comme d'autre part, l'ensemble $\{c_1(t, s); -\infty < t < \infty\}$ est partout dense dans $L_2(-\infty, +\infty)$ (puisque $\mathcal{E}_t(x_1) = \mathcal{E}_t(d\xi_1)$ pour tout t), le produit scalaire $g(t)$, (2.4.44) dans $L_2(-\infty, \infty)$ ne peut s'annuler que si $c_2(t, s)$ est nulle pour presque tout s . Ce fait entraîne que la densité d'innovation de $y'(t)$ doit être nulle presque partout; mais, comme cette densité est analytique, elle est alors identiquement nulle.

Puisqu'il en est ainsi, $y'(t)$ qui vérifie l'hypothèse H (c'est-à-dire $\langle y'(t/s), y'(\tau/s) \rangle$ est analytique en t, τ et s ($s \leq t, \tau$)) est dérivable en m. q. d'après la proposition 5.

Par un raisonnement semblable à celui que nous venons de faire, on peut démontrer que la dérivée seconde possède une densité d'innovation identiquement nulle, et que la dérivée d'ordre trois existe en m. q.; par suite, on peut démontrer que $y(t)$ est indéfiniment dérivable.

PROPOSITION 7. — *Tout processus $y(t)$ satisfaisant à l'hypothèse H [c'est-à-dire si $\langle y(t/s), y(\tau/s) \rangle$ est analytique en t, τ et s ($s \leq t, \tau$)] et indéfiniment dérivable est un processus analytique.*

Puisque toutes les dérivées partielles de la fonction de covariance $\Gamma_y(t, \tau)$ de $y(t)$ sont continues à la traversée de la diagonale principale, le développement

$$\Gamma_y(t+h, t+k) - \Gamma_y(t+h, t) - \Gamma_y(t, t+k) + \Gamma_y(t, t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{h^m \cdot k^n}{m! n!} \left[\frac{\partial^{m+n} \Gamma_y(t, t)}{\partial t^m \partial \tau^n} \right]_{\tau=t}$$

est valable, $\Gamma_y(t, \tau)$ étant analytique au voisinage de l'axe $t = \tau$, $y(t)$ est un processus analytique au sens de la définition 2.4.1.

Nous avons vu que l'existence de la composante $x_1(t)$ ayant les propriétés indiquées plus haut interdit l'existence d'une autre composante non déterministe et que la différence $y(t) = x(t) - x_1(t)$ ne peut qu'être analytique. Nous allons voir qu'il en est toujours ainsi, si $x(t)$ n'est pas indéfiniment dérivable, c'est-à-dire qu'il existe une composante purement non déterministe, qui vérifie l'hypothèse H, que nous représenterons par

$$(2.4.45) \quad u(t) = \int_{-\infty}^t c(t, s) d\xi(s)$$

et que $v(t) = x(t) - u(t)$ est un processus analytique.

PROPOSITION 8. — *Soit n (un entier fini) l'ordre de dérivabilité le plus élevé de $x(t)$, le noyau $c(t, s)$ correspondant à sa partie purement non déterministe est donné par*

$$(2.4.46) \quad c(t, s) = \frac{1}{K_{n+1}(s)} \left[\frac{\partial^{n+1} \Gamma(t, s)}{\partial s^{n+1}} - \langle x(t), \mathbf{x}^{(n+1)}(s/s) \rangle \right], \quad \text{où}$$

$$(2.4.47) \quad K_{n+1}^2(s) = \lim_{\tau \downarrow t} \left[\frac{\partial^{2n+1} \Gamma(t, \tau)}{\partial t^{n+1} \partial \tau^n} - \frac{\partial^{2n+1} \bar{\Gamma}(t, \tau)}{\partial t^n \partial \tau^{n+1}} \right]$$

Si $x^{(n)}(t)$, dérivée n -ième en m. q. d'ordre le plus élevé de $x(t)$ n'est pas dérivable, c'est qu'elle possède une densité d'innovation $K_{n+1}^2(t)$. Cette densité d'innovation est une fonction analytique,

puisque $x^{(n)}(t)$ vérifie l'hypothèse H. L'expression ci-dessus de cette densité n'est que l'application de la formule (2.3.4) au processus $x^{(n)}(t)$.

Nous pouvons définir une fonction aléatoire $\xi(t)$ à accroissements homogènes non corrélés à partir des innovations de $x^{(n)}(t)$, comme nous avons défini $\xi_1(t)$ à partir des innovations de $x(t)$ lorsque $K_1^2(t)$ était supposée non identiquement nulle. Nous désignons par $u(t)$ (2.4.45) la projection de $x(t)$ sur $\mathcal{E}_t(d\xi)$. $c(t, s) \bar{c}(\tau, s)$ étant une fonction analytique de t, τ et s ($s \leq t, \tau$), $u(t)$ est un processus à projections analytiques vérifiant l'hypothèse H. Cette dernière propriété n'est qu'une conséquence directe de l'expression (2.4.46) que nous pouvons retrouver de la manière suivante, sans reprendre entièrement le raisonnement qui nous avait permis de trouver l'expression de $c_1(t, s)$.

$$\begin{aligned} c(t, s) &= \lim_{h \downarrow 0} \frac{1}{h} \langle x(t), \xi(s+h) - \xi(s) \rangle \\ &= \frac{1}{K_{n+1}(s)} \lim_{h \downarrow 0} \frac{1}{h} \langle x(t), x^{(n)}(s+h) - x^{(n)}(s) \rangle \\ &= \frac{1}{K_{n+1}(s)} \left[\lim_{h \downarrow 0} \frac{1}{h} \langle x(t), x^{(n)}(s+h) - x^{(n)}(s) \rangle - \lim_{h \downarrow 0} \frac{1}{h} \langle x(t), x^{(n)}(s+h/s) - x^{(n)}(s) \rangle \right] \end{aligned}$$

L'unicité de la composante non déterministe de $x(t)$ se démontre en appliquant les considérations de la proposition 6.

1.2.5. Conclusions : justification des hypothèses ; une représentation intégrale duale ; cas des processus aléatoires stationnaires du second ordre.

a) *Justification de l'hypothèse H.* — Considérons un processus aléatoire à projections dérivables, ayant pour représentation canonique

$$(2.5.1) \quad x(t) = \int_{-\infty}^t c(t, s) d\xi(s)$$

$\xi(s)$ étant à accroissements homogènes non corrélés. Supposons que $c(t, s)$ soit continu par rapport à t et à s ($t \geq s$) et dérivable par rapport à t , quel que soit s , et que $\mathbf{x}'(t/t)$ puisse s'écrire sous la forme

$$\mathbf{x}'(t/t) = \int_{-\infty}^t \frac{\partial}{\partial t} c(t, s) d\xi(s).$$

Supposons aussi, que $\frac{\partial}{\partial t} c(t, s)$ soit continu par rapport à t et à s ($t \geq s$). Alors, nous avons pour $\tau > t$

$$(2.5.2) \quad \frac{\partial}{\partial t} \Gamma(t, \tau) = c(t, t) \bar{c}(\tau, t) + \int_{-\infty}^t \frac{\partial c(t, s)}{\partial t} \cdot \bar{c}(\tau, s) ds$$

$$(2.5.3) \quad \frac{\partial}{\partial \tau} \Gamma(t, \tau) = \int_{-\infty}^t c(t, s) \cdot \frac{\partial \bar{c}(\tau, s)}{\partial \tau} ds$$

$$(2.5.4) \quad \frac{\partial^2}{\partial t \partial \tau} \Gamma(t, \tau) = c(t, t) \frac{\partial}{\partial \tau} \bar{c}(\tau, t) + \int_{-\infty}^t \frac{\partial c(t, s)}{\partial t} \cdot \frac{\partial \bar{c}(\tau, s)}{\partial \tau} ds$$

Il est clair qu'un tel processus satisfait aux conditions du théorème 2.3.1. Par conséquent, nous trouvons l'expression de la densité d'innovation

$$K_1^2(t) = \lim_{\tau \downarrow t} \left[\frac{\partial}{\partial t} \Gamma(t, \tau) - \frac{\partial}{\partial \tau} \bar{\Gamma}(t, \tau) \right]$$

que nous supposons non identiquement nulle. Mais de plus, nous tirons de (2.5.2)

$$\begin{aligned} c(\tau, t) &= \frac{1}{c(t, t)} \left[\frac{\partial}{\partial t} \bar{\Gamma}(t, \tau) - \int_{-\infty}^t \frac{\partial \bar{c}(t, s)}{\partial t} \cdot c(\tau, s) ds \right] \\ &= \frac{1}{c(t, t)} \left[\frac{\partial}{\partial t} \Gamma(\tau, t) - \langle x(\tau), \mathbf{x}'(t/t) \rangle \right]. \end{aligned}$$

ce qui n'est autre que l'expression (2.4.24), que nous avons établie sous d'autres conditions.

Si, par contre, $K_1^2(t) = |c(t, t)|^2 \equiv 0$, $c(t, s)$ reste indéterminé. Les relations (2.5.2, 3 et 4) montrent que $x(t)$ est faiblement dérivable et que cette dérivée faible coïncide p. s. avec $\mathbf{x}'(t/t)$. Supposons maintenant, que $\mathbf{x}''(t/t)$ existe et puisse être représentée par

$$\mathbf{x}''(t/t) = \int_{-\infty}^t \frac{\partial^2}{\partial t^2} c(t, s) d\xi(s),$$

$\frac{\partial^2}{\partial t^2} c(t, s)$ étant continu par rapport à t et à s ($s \leq t$). On peut démontrer la dérivabilité en m. q. de $x(t)$ en remarquant que

$$\lim_{h \downarrow 0} \frac{1}{h} \|x(t+h) - x(t)\|^2 = \left[\frac{\partial^2 \Gamma(t, \tau)}{\partial t \partial \tau} \right]_{\tau=t} = \|\mathbf{x}'(t/t)\|^2.$$

Si, de plus, $x'(t)$ satisfait aux mêmes conditions que le processus $x(t)$ (2.5.1), en supposant que sa densité d'innovation

$$K_2^2(t) = \lim_{s \downarrow t} \left| \frac{\partial}{\partial t} c(t, s) \right|^2$$

n'est pas identiquement nulle, nous pouvons trouver l'expression du noyau $c(t, s)$, en calculant à partir de (2.5.2) la dérivée $\frac{\partial^2 \Gamma(t, \tau)}{\partial t^2}$. Ce qui donne

$$c(\tau, t) = \frac{1}{\left(\frac{\partial}{\partial t} \bar{c}(t, s) \right)_{s=t}} \left[\frac{\partial^2 \Gamma(\tau, t)}{\partial t^2} - \langle x(\tau), \mathbf{x}''(t/t) \rangle \right].$$

Et en continuant de cette façon nous retrouverions l'expression (2.4.10) du noyau $c(t, s)$ dans le cas où l'ordre le plus élevé de dérivabilité de $x(t)$ ne dépasse pas un nombre fini n .

Par conséquent, si l'on se borne à ce mode de construction,

1° l'introduction des processus à projections indéfiniment dérivables (ou bien à projections analytiques, pour satisfaire à certaines conditions de régularité),

2° l'existence et la continuité des dérivées

$$\frac{\partial^{m+n} \Gamma(t, \tau)}{\partial t^m \partial \tau^n} \quad \text{pour } t \neq \tau,$$

3° la continuité des limites
$$\lim_{h \downarrow 0} \frac{\partial^{m+n} \Gamma(t, \tau)}{\partial t^m \partial \tau^n},$$

4° la dérivabilité en m. q. lorsque la densité d'innovation est identiquement nulle, s'imposent d'elles-mêmes.

Nous avons vu que les conditions du théorème 2.3.1 entraînent l'existence d'une composante $x_1(t)$ possédant une représentation canonique (2.3.9); mais nous allons voir, d'après le théorème suivant, que le problème reste entier en ce qui concerne la représentation de $y(t) = x(t) - x_1(t)$. De toutes les manières, si $x(t)$ possède les propriétés 1°, 2°, ..., 4° citées ci-dessus, il apparaît que $x_1(t)$ aussi doit les posséder, pour que la méthode que nous avons préconisée puisse être applicable à $y(t)$.

THÉORÈME 2.5.1. — Dans les conditions du théorème 2.3.1, la projection

$$x_1(t) = \int_{-\infty}^t c_1(t, s) d\xi_1(s)$$

de $x(t)$ sur $\mathcal{E}_t(d\xi_1)$ satisfait aussi aux conditions du théorème 2.3.1, le noyau $c_1(t, s)$ étant donné par

$$c_1(t, s) = \frac{1}{K_1(s)} \left[\frac{\partial \Gamma(t, s)}{\partial s} - \langle x(t), \mathbf{x}'(s/s) \rangle \right]$$

pour tout t et pour presque tous s . Si de plus, on suppose $K_1(t) \neq 0$, le processus $y(t) = x(t) - x_1(t)$ admet une densité d'innovation identiquement nulle et dans le cas où $\frac{\partial^2 \Gamma(t, \tau)}{\partial t \partial \tau}$ existe et est continue, $y(t)$ est faiblement dérivable.

Démonstration. — Nous reprenons la relation (2.2.4), reliant $Z_1(t)$, $x(t)$, et $x'(t/t)$, que nous avons établie par le théorème 2.2.1.

$$\int_{t_0}^t dZ_1(s) = x(t) - x(t_0) - \int_{t_0}^t \mathbf{x}'(s/s) ds.$$

Nous avons, compte tenu de la continuité absolue de $F_1(t)$,

$$K_1(s) c_1(t, s) = \frac{\partial \Gamma(t, s)}{\partial s} - \langle x(t), \mathbf{x}'(s/s) \rangle$$

pour tout s et t ; expression continue par rapport à t et à s ($s \leq t$). Ce qui fait que $c_1(t, s)$ est bien défini toutes les fois que $K_1(s) \neq 0$. $c_1(t, s)$ peut alors être défini par cette expression pour presque tout s , si l'on convient de prendre $c_1(t, s) = 0$ toutes les fois que $K_1(s) = 0$. Nous supposons que $K_1(t)$ ne s'annule jamais; auquel cas $c_1(t, s)$ est continu par rapport à s et à t ($s \leq t$).

Il est évident que, comme $x(t)$, $x_1(t)$ est continu en m. q. et à projections dérivables et $\mathbf{x}'_1(t/t)$ est aussi continu en m. q. Alors si $\Gamma_1(t, \tau)$ désigne la covariance de $x_1(t)$, la dérivée partielle

$$(2.5.5) \quad \frac{\partial \Gamma_1(t, \tau)}{\partial \tau} = \langle x_1(t), \mathbf{x}'_1(\tau/t) \rangle$$

est continue pour $t \leq \tau$; d'autre part nous avons

$$\frac{\partial \Gamma_1(t, \tau)}{\partial t} = \lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} c_1(t+h, s) \bar{c}_1(t, s) ds + \langle \mathbf{x}'(t/t), x_1(\tau) \rangle$$

$$\text{Comme } \lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} c_1(t+h, s) \bar{c}_1(\tau, s) ds = \bar{c}_1(t, t) c_1(\tau, t) = \frac{\partial \bar{\Gamma}(\tau, t)}{\partial t} = \langle \mathbf{x}'(t/t), x(\tau) \rangle$$

est défini pour tout t ,

$$(2.5.6) \quad \frac{\partial \Gamma_1(t, \tau)}{\partial t} = \left[-\frac{\partial \Gamma(t, \tau)}{\partial t} - \langle \mathbf{x}'(t/t), x(\tau) \rangle \right] + \langle \mathbf{x}'_1(t/t), x_1(\tau) \rangle$$

existe et est continue pour tout $t \leq \tau$.

L'existence et la continuité de $\frac{\partial^2 \Gamma_1(t, \tau)}{\partial t^2}$ peuvent être démontrées de la même manière et nous avons

$$(2.5.7) \quad \frac{\partial^2 \Gamma_1(t, \tau)}{\partial t \partial \tau} = \left[\frac{\partial^2 \Gamma(t, \tau)}{\partial t \partial \tau} - \langle \mathbf{x}'(t/t), \mathbf{x}'(\tau/\tau) \rangle \right] + \langle \mathbf{x}'_1(t/t), \mathbf{x}'_1(\tau/\tau) \rangle ;$$

x_1 satisfait donc aux conditions du théorème 2.3.1 et sa densité d'innovation est la même que celle de $x(t)$, puisque

$$\lim_{\tau \downarrow t} \left[\frac{\partial \Gamma_1(t, \tau)}{\partial t} - \frac{\partial \bar{\Gamma}_1(t, \tau)}{\partial \tau} \right] = \lim_{\tau \downarrow t} \left[\frac{\partial \Gamma(t, \tau)}{\partial t} - \frac{\partial \bar{\Gamma}(t, \tau)}{\partial \tau} \right] = K_1^2(t).$$

Les expressions (2.5.5, 6 et 7) montrent que, si $\Gamma_y(t, \tau)$ désigne la fonction de covariance de $y(t) = x(t) - x_1(t)$

$$\frac{\partial \Gamma_y(t, \tau)}{\partial t} = \langle \mathbf{y}'(t/t), y(\tau) \rangle$$

$$\frac{\partial^2 \Gamma_y(t, \tau)}{\partial t \partial \tau} = \langle \mathbf{y}'(t/t), \mathbf{y}'(\tau/\tau) \rangle$$

quels que soient t et τ . Ce qui entraîne la dérivabilité faible de $y(t)$.

Que faut-il supposer de plus sur $\Gamma_y(t, \tau)$ pour que $y(t)$ soit dérivable en m. q.? Il suffit, pour cela, que l'on ait

$$\lim_{h \downarrow 0} \left\| \frac{y(t+h) - y(t)}{h} \right\|^2 = \left[\frac{\partial^2 \Gamma_y(t, \tau)}{\partial t \partial \tau} \right]_{\tau=t} = \|\mathbf{y}'(t/t)\|^2$$

et cette condition est vérifiée si $\frac{\partial^2 \Gamma_y(t, \tau)}{\partial t^2}$ est une fonction continue, quels que soient t et τ . Mais si

l'on veut voir par quelle hypothèse sur $x(t)$, l'existence et la continuité de $\frac{\partial^2 \Gamma_y(t, \tau)}{\partial t^2}$ sont assurées, on s'aperçoit à l'examen de la proposition 5 que l'on ne peut pas éviter de supposer l'existence de $\mathbf{x}''(t/t)$ et la dérivabilité de $\|\mathbf{x}(t/s)\|^2$ et de $\|\mathbf{x}_1(t/s)\|^2$ par rapport à s , conditions qui sont contenues dans l'hypothèse H. On voit ainsi, l'hypothèse H se justifier d'elle-même, si de plus $y(t)$ possède une dérivée en m. q. d'ordre très élevé.

Remarque. — Considérons la relation (2.2.4)

$$(2.5.8) \quad \int_{t_0}^t K_1(s) d\xi_1(s) = x(t) - x(t_0) - \int_{t_0}^t \mathbf{x}'(s/s) ds$$

qui peut aussi être exprimée par l'équation aux accroissements du premier ordre

$$(2.5.9) \quad dx(t) = K_1(t) d\xi_1(t) + \mathbf{x}'(t/t) dt$$

qui a servi de point de départ à P. Lévy ([1], [2], [3]) pour l'étude des processus aléatoires admettant une représentation canonique. Il ressort du théorème 2.5.1 que la solution de l'équation (2.5.8) ou de (2.5.9) satisfaisant aux conditions du théorème 2.3.1 n'est pas unique. Il existe une solution donnée par

$$x_1(t) = \int_{-\infty}^t c_1(t, s) d\xi_1(s)$$

ayant pour densité d'innovation $K_1^{\pm}(t)$ et toutes les autres solutions peuvent se mettre sous la forme

$$x(t) = x_1(t) + y(t)$$

où $y(t) [_ \mathcal{E}_t(s_1)]$ possède une densité d'innovation identiquement nulle. Il est évident que la classe de tels processus aléatoires est encore trop riche, pour qu'il soit possible de donner une méthode de construction sans être obligé d'introduire des conditions plus restrictives que celle du théorème 2.3.1. Comme nous l'avons vu, c'est seulement par l'hypothèse H que nous sommes arrivés à supprimer cette incertitude sur $y(t)$, tout en gardant pour base de construction le théorème 2.3.1.

b) *Une représentation intégrale de (2.4.8) par dualité.* — Si l'on convient de prendre $\xi(0) = 0$, $\xi(t)$ peut être représenté par

$$(2.5.10) \quad \xi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{e^{i\omega t} - 1}{i\omega} d\xi^*(\omega)$$

où $\xi^*(\omega)$ est aussi une fonction à accroissements homogènes non corrélés. Réciproquement, l'inversion de l'intégrale (2.5.10) donne

$$(2.5.11) \quad \xi^*(\omega) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \frac{e^{-i\omega t} - 1}{-it} d\xi(t).$$

$u(t)$ peut être mis sous la forme

$$(2.5.12) \quad u(t) = \int_0^{+\infty} c(t, t-\tau) d\xi(t-\tau).$$

La fonction $c(t, t-\tau)$ étant à carré du module sommable en tant que fonction de τ , elle possède une transformée de Fourier-Plancherel

$$(2.5.13) \quad \varphi(t, \omega) = \int_0^{+\infty} e^{-i\tau} c(t, t-\tau) d\tau,$$

qui est une fonction à carré du module sommable par rapport à ω et pour les valeurs complexes de ω , elle est holomorphe dans le demi-plan défini par $\text{Im } \omega < 0$. (Remarquer l'analogie avec la fonction $\Gamma_u(z)$ introduite dans le § 1.2). Compte tenu de (2.5.10), l'application du théorème de Parseval à l'intégrale (2.5.12) donne

$$(2.5.14) \quad u(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t} \varphi(t, \omega) d\xi^*(\omega).$$

La fonction de covariance de $u(t)$ est évidemment donnée par

$$(2.5.15) \quad \Gamma_u(t, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega(t-\tau)} \varphi(t, \omega) \overline{\varphi(\tau, \omega)} d\omega.$$

En ce qui concerne la représentation spectrale de $v(t)$, nous remarquerons d'abord, qu'un processus analytique est entièrement déterminé par la donnée d'un nombre, tout au plus dénombrable, de ses valeurs. Par conséquent, la puissance $\mathcal{E}(v)$ est tout au plus égale à la puissance des entiers positifs. Nous pouvons donc appliquer la méthode du § 1.2 pour la représentation de $v(t)$. Alors, $v(t)$ peut être représenté par

$$(2.5.16) \quad v(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t} \psi(t, \omega) dZ_w(\omega)$$

où $Z_w(\omega)$ est la fonction spectrale d'un processus stationnaire déterministe, c'est-à-dire, une fonction aléatoire à accroissements non corrélés, dont les accroissements ne diffèrent de zéro que sur un ensemble de mesure de Lebesgue nulle.

Si maintenant, on introduit la fonction spectrale

$$(2.5.17) \quad Z(\omega) - Z(\omega_1) = \int_{\omega_1}^{\omega} d\xi^*(\omega) + dZ_w(\omega)$$

on peut représenter $x(t)$ par

$$(2.5.18) \quad x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} \varphi^*(t, \omega) dZ(\omega), \text{ où}$$

$\varphi^*(t, \omega) = \varphi(t, \omega)$ pour presque tout ω (mesure de Lebesgue)

$\varphi^*(t, \omega) = \psi(t, \omega)$ pour presque tout ω ($F_w = \|Z_w\|^2$ — mesure).

Par conséquent, le théorème 1.2.3 reste vrai pour un processus permanent vérifiant l'hypothèse H.

c) *Cas des processus stationnaires du second ordre.* — Tout processus $x(t)$ stationnaire du second ordre continu en m. q. peut être représenté par

$$(2.5.19) \quad x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} dX(\omega) \quad (\text{p. s.}),$$

où $X(\omega)$ est un processus à accroissements non corrélés. $x(t)$ admet une décomposition orthogonale.

$$(2.5.20) \quad x(t) = u(t) + v(t)$$

$u(t)$, étant purement non déterministe et $v(t)$ déterministe et tout processus stationnaire du second ordre purement non déterministe, admet la représentation canonique propre

$$(2.5.21) \quad u(t) = \int_{-\infty}^t c(t-s) d\xi(s)$$

où $\xi(t)$ est à accroissements homogènes non corrélés et $c(t) \in L_2(0, \infty)$. Soit $\varphi(\omega)$ la transformée de Fourier-Plancherel de $c(t)$. $\varphi(\omega)$ est holomorphe dans le demi-plan $\text{Im } \omega < 0$ et n'y a pas de zéros. $F(\omega) = \|X(\omega)\|^2$ possède une dérivée presque partout égale à $|\varphi(\omega)|^2$ et cette dérivée s'appelle densité spectrale.

THÉORÈME 2.5.2. — *Tout processus stationnaire du second ordre à densité spectrale rationnelle vérifie l'hypothèse H, c'est-à-dire, si $x(t)$ désigne le processus, $\langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle$ est une fonction analytique de t, τ et s ($s \leq t, \tau$).*

Démonstration. — Posons

$$(2.5.22) \quad |\varphi(\omega)|^2 = |c|^2 \frac{\prod_{j=1}^{2M} (\omega - w_j)}{\prod_{j=1}^{2N} (\omega - z_j)}, \quad w_j' \neq z_k' \quad (j, k = 1, 2, \dots), \quad M < N$$

Comme $|\varphi|^2$ est intégrable, le dénominateur n'a pas de racines réelles. De plus, les zéros du numérateur et du dénominateur, qui ne sont pas les mêmes, sont deux à deux imaginaires conjugués. $|\varphi(\omega)|^2$ peut alors s'écrire :

$$(2.5.23) \quad |\varphi(\omega)|^2 = \left| c \frac{\prod_{j=1}^M (\omega - w_j)}{\prod_{j=1}^N (\omega - z_j)} \right|^2 \quad \text{Im } w_j \geq 0, \quad \text{Im } z_j > 0.$$

Puisque $\varphi(\omega)$ doit être holomorphe et ne doit pas avoir de zéros dans le demi-plan $\text{Im } \omega < 0$, nous pouvons prendre

$$(2.5.24) \quad \varphi(\omega) = c \cdot \frac{\prod_{j=1}^M (\omega - w_j)}{\prod_{j=1}^N (\omega - z_j)}$$

$x(t)$ aura par conséquent la représentation

$$(2.5.25) \quad x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t} \frac{\prod_{j=1}^M (\omega - w_j)}{\prod_{j=1}^N (\omega - z_j)} \cdot d\xi^*(\omega).$$

Le noyau $c(t)$ correspondant à la représentation (2.5.21) peut être mis sous la forme

$$(2.5.26) \quad c(t) = \prod_{j=1}^N c_j(t) e^{itz_j}$$

où les fonctions $c_j(t)$ sont des polynômes. On peut aisément montrer que $c(t)$ est analytique pour $t \geq 0$ et ses dérivées appartiennent toutes à $L_2(0, \infty)$. Donc,

$$(2.5.27) \quad \mathbf{x}(t/s) = \int_{-\infty}^s \prod_{j=1}^N c_j(t-r) e^{i(t-r)z_j} d\xi(r)$$

est analytique en t ($t \geq s$) et d'autre part

$$(2.5.28) \quad \langle \mathbf{x}(t/s), \mathbf{x}(\tau/s) \rangle = \int_{-\infty}^s \prod_{j=1}^N c_j(t-r) e^{i(t-r)z_j} \times \prod_{j=1}^N \bar{c}_j(\tau-r) e^{-i(\tau-r)\bar{z}_j} dr$$

est une fonction analytique de t , τ et de s ($s \leq t, \tau$).

Remarquons maintenant, que les dérivées de $c(t)$ jusqu'à l'ordre $(N-M-2)$ (inclus) sont toutes continues au point $t=0$, donc nulles. Si l'on désigne par $c^{(k)}(0-)$ la dérivée k -ième à gauche du point $t=0$ et par $c^{(k)}(0+)$, la dérivée à droite, la quantité

$$\frac{1}{2\pi i} [c^{(k)}(0+) - c^{(k)}(0-)]$$

est égale au résidu à l'infini de la fonction $\frac{1}{2\pi} (i\omega)^k \varphi(\omega)$. Remarquant que $c^{(k)}(0-) = 0$ quel que soit k , nous avons

$$\begin{cases} c^{(k)}(0+) = 0, & \text{pour } k \leq N-M-2 \\ c^{(k)}(0+) = -c^{(N-M-k)}(0-), & \text{pour } k = N-M-1 \end{cases}$$

et comme toutes les dérivées de $c(t)$ sont continues pour $t > 0$, la première dérivée qui présente une discontinuité à l'origine est la dérivée $(N-M-1)$ -ième. Nous voyons que $x(t)$ est dérivable en m. q. jusqu'à l'ordre $(N-M-1)$ (inclus). La densité d'innovation de la dérivée $N-M-1$ est donnée par

$$(2.5.29) \quad \lim_{h \downarrow 0} \frac{1}{h} \|x^{(k)}(t+h) - x^{(k)}(t+h/t)\|^2 = |c|^2, \quad k = N-M-1.$$

On peut voir, de la même manière que pour la fonction $c(t)$, que la fonction de corrélation $\Gamma(t)$ est analytique pour $t \neq 0$ et

$$\begin{aligned} \Gamma^{(2k+1)}(0+) - \Gamma^{(2k+1)}(0-) &= 0 & \text{pour } k = N-M-2 \\ \Gamma^{(2k+1)}(0+) - \Gamma^{(2k+1)}(0-) &= |c|^2 i^{2N-2M} & \text{pour } k = N-M-1. \end{aligned}$$

Nous avons donc

$$(2.5.30) \quad \begin{aligned} &\lim_{h \downarrow 0} \frac{1}{h} \|x^{(k)}(t+h) - x^{(k)}(t+h/t)\|^2 \\ &= (-1)^{k-1} [\Gamma^{(2k+1)}(0+) - \Gamma^{(2k+1)}(0-)] = |c|^2 \quad \text{pour } k = N-M-1, \end{aligned}$$

ce qui est conforme au résultat (2.5.29).

CHAPITRE II : PRÉVISION, FILTRAGE ET INFORMATION.

2.1. DÉBITS D'ENTROPIE ET D'INFORMATION.

Soit $(X, \mathfrak{A}, P)$ un espace de probabilité (X étant un espace euclidien, $\mathfrak{A}$ la σ -algèbre des ensembles boréliens dans X , et P une mesure de probabilité sur $\mathfrak{A}$). Si la mesure P possède une densité $p(x)$, l'entropie différentielle de la v. a. prenant ses valeurs dans X est définie par

$$H(x) = - \int p(x) \log p(x) dx.$$

Considérons un espace de probabilité $(X \times Y, \mathfrak{A} \times \mathfrak{B}, \omega)$, où l'espace mesurable $(X \times Y, \mathfrak{A} \times \mathfrak{B})$ est formé par le produit cartésien de deux espaces mesurables $(X, \mathfrak{A})$ et $(Y, \mathfrak{B})$ définis comme ci-dessus. Soient P et Q les mesures induites par ω , respectivement sur $\mathfrak{A}$ et $\mathfrak{B}$. Supposons que $\omega \ll P \times Q$

et désignons par $p(x, y)$ la densité Radon-Nikodym correspondante. Si $P(\cdot/y)$ est la probabilité conditionnelle définie sur $\mathfrak{X}$, elle est mesurable par rapport à $\mathfrak{Y}$ comme fonction de y et on a pour $A \in \mathfrak{X}$ et $B \in \mathfrak{Y}$

$$\omega(A \times B) = \int_A \int_B p(x, y) dP(x) dQ(y) = \int_B P(A/y) dQ(y).$$

De sorte que, la version de $P(\cdot/y)$, définie par

$$P(A/y) = \int_A p(x, y) dP(x), \quad A \in \mathfrak{X}$$

presque sûrement Q -mesure, est une mesure de probabilité absolument continue par rapport à la mesure P .

On peut définir la même propriété, pour la probabilité conditionnelle $P(\cdot/x)$ avec l'hypothèse $\omega \ll P \times Q$. On appelle quantité d'information mutuelle entre deux v. a. $x \in X$ et $y \in Y$, la quantité

$$I(x, y) = \int \log p(x, y) d\omega = \int p(x, y) \log p(x, y) dP.dQ.$$

Si la mesure $P(\cdot/y)$ admet une densité $p(y/x)$, on appelle entropie (différentielle) conditionnelle de x connaissant y , la quantité

$$H(x/y) = - \int \log p(x/y) d\omega.$$

On peut facilement vérifier, dans le cas où les densités $p(x, y)$, $p(x)$ et $p(y)$ existent, que

$$I(x, y) = H(x) - H(x/y) = H(y) - H(y/x) = H(x) + H(y) - H(x, y).$$

Tout espace de probabilité $(X, \mathfrak{X}, P)$ est appelé une source d'information. Soit un espace mesurable

$$(X, \mathfrak{X}) = \prod_{i=-\infty}^{\infty} (X_i, \mathfrak{X}_i),$$

formé du produit cartésien infini des espaces mesurables $(X_i, \mathfrak{X}_i)$, sur lequel nous définissons une mesure de probabilité P . Nous identifierons l'indice i au temps. La suite d'éléments $\{x_i\}$ sera un processus aléatoire caractérisant la source d'information.

Nous supposons que la mesure P possède une densité. Soit $p(x_{-n}, \dots, x_0, \dots, x_n)$ la densité de probabilité correspondant aux variables aléatoires $x_{-n}, \dots, x_0, \dots, x_n$. Nous définirons le débit d'entropie de la source $(X, \mathfrak{X}, P)$ par l'expression

$$\bar{H}(x) = \lim_{n \rightarrow \infty} - \frac{1}{2n+1} \int \dots \int p(x_{-n}, \dots, x_n) \log p(x_{-n}, \dots, x_n) dx_{-n}, \dots, dx_n$$

si elle existe.

Supposons maintenant, que tous les $(X_i, \mathfrak{X}_i)$ s'identifient à $(X_0, \mathfrak{X}_0)$ et que la mesure de probabilité P sur $(X, \mathfrak{X})$ soit stationnaire. Nous désignerons la source ainsi définie par $(X^1, \mathfrak{X}^1, P)$. Nous avons, d'après la stationnarité de la mesure P ,

$$H(x_{-n}, \dots, x_0, \dots, x_n) = H(x_{-2n}, \dots, x_{-n}, \dots, x_0).$$

Le débit d'entropie de la source stationnaire pourra être calculé par

$$\bar{H}(x) = \lim_{n \rightarrow \infty} \frac{1}{n+1} H(x_{-n}, \dots, x_0).$$

D'autre part, l'entropie conditionnelle

$$H(x_0/x_{-1}, \dots, x_{-n}) = - \int \dots \int p(x_{-n}, \dots, x_0) \log \frac{p(x_{-n}, \dots, x_0)}{p(x_{-n}, \dots, x_{-1})} dx_{-n}, \dots, dx_0$$

pouvant s'écrire de la manière suivante

$$H(x_0/x_{-1}, \dots, x_{-n}) = H(x_{-n}, \dots, x_0) - H(x_{-n}, \dots, x_{-1}),$$

nous avons, compte tenu de la stationnarité de la mesure

$$H(x_{-n}, \dots, x_0) = H(x_0/x_{-1}, \dots, x_{-n}) + H(x_{-1}/x_{-2}, \dots, x_{-n}) + \dots + H(x_{-n}) = \sum_{j=0}^n H(x_0/x_{-1}, \dots, x_{-j})$$

où nous avons posé $H(x_0/x_0) = H(x_0)$. Désignons par H_j , l'entropie conditionnelle $H_j = H(x_0/x_{-1}, \dots, x_{-j})$.

THÉOREME A. — Si H_j possède une limite lorsque $j \rightarrow \infty$, le débit d'entropie de la source stationnaire $(X^I, \mathfrak{X}^I, P)$ existe et il est égal à cette limite.

Démonstration. — Etant donné un nombre $\varepsilon > 0$ arbitrairement petit, il existe un entier j_0 tel que si $j > j_0$

$$|H_j - \bar{H}| < \varepsilon/2, \quad \text{où } \bar{H} = \lim_{j \rightarrow \infty} H_j.$$

Soit un entier $j_1 > j_0$, tel que $\frac{1}{j_1} \sum_{j=0}^{j_0} |H_j - \bar{H}| < \varepsilon/2$. Si $n > j_1$, nous avons

$$\begin{aligned} & \left| \left(\frac{1}{n+1} \sum_{j=0}^n H_j \right) - \bar{H} \right| = \left| \frac{1}{n+1} \sum_{j=0}^n (H_j - \bar{H}) \right| \\ & \leq \left| \frac{1}{n+1} \sum_{j=0}^{j_0} (H_j - \bar{H}) \right| + \left| \frac{1}{n+1} \sum_{j=j_0+1}^n (H_j - \bar{H}) \right| < \frac{1}{j_1} \sum_{j=0}^{j_0} |H_j - \bar{H}| + \frac{n - j_0}{n+1} \cdot \frac{\varepsilon}{2} < \varepsilon \end{aligned}$$

Nous avons donc

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} H(x_0, \dots, x_n) = \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{j=0}^n H(x_0/x_1, \dots, x_j) = \lim_{n \rightarrow \infty} H(x_0/x_1, \dots, x_j) = \bar{H}.$$

Soit maintenant un espace de probabilité $(X \times Y, \mathfrak{X} \times \mathfrak{Y}, \omega)$, où

$$(X \times Y, \mathfrak{X} \times \mathfrak{Y}) = \prod_{i=-\infty}^{\infty} (X_i \times Y_i, \mathfrak{X}_i \times \mathfrak{Y}_i) = \prod_{i=-\infty}^{+\infty} (X_i, \mathfrak{X}_i) \times \prod_{i=-\infty}^{\infty} (Y_i, \mathfrak{Y}_i)$$

Désignons par P_n et Q_n les mesures induites par ω respectivement dans $\prod_{i=-n}^n \mathfrak{X}_i$ et $\prod_{i=-n}^n \mathfrak{Y}_i$ et supposons que $\omega \ll P_n \times Q_n$ quel que soit n . Soit $g_n(x, y)$ la densité de Radon-Nikodym correspondante. Nous appellerons débit d'information mutuelle entre les deux sources $(X, \mathfrak{X}, P)$ et $(Y, \mathfrak{Y}, Q)$ la quantité

$$I(x, y) = \lim_{n \rightarrow \infty} \frac{1}{2n+1} \int \log g_n(x, y) d\omega = \lim_{n \rightarrow \infty} \frac{1}{2n+1} \int g_n(x, y) \log g_n(x, y) dP dQ.$$

Remarquons que, dans le cas stationnaire, la limite $I(x, y)$ peut être calculée comme le débit d'information entre les sources $(\prod_{i=-\infty}^0 X_i, \mathfrak{X}_i, \bar{P})$ et $(\prod_{i=-\infty}^0 Y_i, \mathfrak{Y}_i, \bar{Q})$ où $\bar{P}$ et $\bar{Q}$ sont respectivement les mesures induites par ω dans $\prod_{i=-\infty}^0 \mathfrak{X}_i$ et $\prod_{i=-\infty}^0 \mathfrak{Y}_i$.

2.2. PRÉVISION LINÉAIRE DES PROCESSUS VECTORIELS ET STATIONNAIRES DU SECOND ORDRE.

(Dans ce qui suit, nous supposons que toutes les v. a. considérées ont une valeur moyenne nulle. Nous désignerons par (A, B) le gramien de deux vecteurs aléatoires A et B ; par Δ , le déterminant; par $(\)^*$, l'adjointe, et par $\text{Tr}(\)$, la trace d'une matrice.)

Soit $\{X_n\}$, une suite de vecteurs aléatoires, chaque élément X_n de la suite ayant q composantes $x_n^1, x_n^2, \dots, x_n^q$ telles que $E |x_n^j|^2 < \infty$, $1 \leq j \leq q$ et quel que soit n .

Dans l'espace de Hilbert $L(X)$ engendré par les éléments de la suite, le produit scalaire de deux éléments X et Y est donné par

$$(2.1) \quad ((X, Y)) = \sum_{j=1}^q E(x^j, \bar{y}^j) = \text{Tr}(X, Y)$$

Alors, la norme d'un élément X est

$$(2.2) \quad \|X\| = \sqrt{((X, X))} = \sqrt{\sum_{j=1}^q E |x^j|^2} = \sqrt{\text{Tr}(X, X)}$$

et cette norme induit une topologie dans $L(X)$.

$L(X)$ a les propriétés suivantes :

- X et Y sont des vecteurs orthogonaux, si et seulement si $(X, Y) = 0$.
- Une suite de vecteurs aléatoires $\{X_n\}$ est une suite orthonormale, si et seulement si $(X_n, X_m) = \delta_{n,m}I$, où $\delta_{n,m} = 1$ si $n = m$ et $= 0$ si $n \neq m$ et I est la matrice unité.
- Un ensemble linéaire $\mathcal{E}$ est un sous-espace dans L , s'il est fermé dans la topologie induite par la norme définie plus haut.
- Si $\mathcal{E}$ est un sous-espace de L , il existe pour tout $X \in L$, un élément $Z \in \mathcal{E}$, tel que

$$\|X - Z\| \leq \|X - Y\| \quad \text{pour tout } Y \in \mathcal{E}$$

Alors $X - Z \perp \mathcal{E}$ et Z s'appelle la projection de X sur $\mathcal{E}$.

Supposons maintenant, que la suite $\{X_n\}$ soit stationnaire au sens large. Il lui correspond une fonction $X(\lambda)$ vectorielle, à q dimensions, à accroissements non corrélés, c'est-à-dire telle que

$$(dX(\lambda), dX(\mu)) = 0 \quad \text{pour } \lambda \neq \mu, \text{ et que}$$

$$(2.3) \quad X_n = \int_{-\pi}^{\pi} e^{in\lambda} dX(\lambda).$$

Nous désignerons par

$$F_{xx}(\lambda) = (X(\lambda), X(\lambda))$$

la fonction de distribution spectrale du processus $\{X_n\}$. $F_{xx}(\lambda)$ est une fonction matricielle d'ordre q , hermitienne et non-décroissante ; en ce sens que, la différence $F_{xx}(\lambda') - F_{xx}(\lambda)$ est une matrice hermitienne non négative pour $\lambda' > \lambda$. La quantité

$$(2.4) \quad \Gamma_{n-m} = (X_n, X_m) = \int_{-\pi}^{\pi} e^{i(n-m)\lambda} dF_{xx}(\lambda) = \Gamma_{n-m}^*$$

est la matrice de covariance des vecteurs X_n et X_m .

Désignons par $\mathbf{X}_n$ la projection de X_n sur le sous-espace linéaire engendré par les éléments X_{n-j} ($j \geq 1$) du processus. La différence $X_n - \mathbf{X}_n$ s'appelle l'innovation du processus à l'instant n . On dit que le processus est à rang complet (full rank) lorsque le gramien de l'innovation a un déterminant non nul.

Si un processus stationnaire $\{X_n\}$ est à rang complet et purement non déterministe, le processus admet p. s. le développement orthogonal suivant :

$$(2.5) \quad X_n = \sum_{m=0}^{\infty} C_m \xi_{n-m}$$

où la suite $\{\xi_n\}$ est orthonormale et $C_0 \xi_n$ est l'innovation à l'instant n , avec

$$(2.6) \quad C_0^2 = G = (X_n - \mathbf{X}_n, X_n - \mathbf{X}_n).$$

A partir du développement orthogonal (2.5), on peut remonter à la fonction génératrice du processus

$$(2.7) \quad \Phi_{xx}(\lambda) = \sum_{m=0}^{\infty} C_m e^{-im\lambda} \quad \text{et démontrer que}$$

$$(2.8) \quad F'_{x,x}(\lambda) = f_{xx}(\lambda) = \Phi_{xx}(\lambda) \cdot \Phi_{xx}^*(\lambda)$$

où $F'_{xx}(\lambda) = f_{xx}(\lambda)$ est la densité spectrale presque partout définie (cf. Wiener et Masani). Les propositions suivantes seront utilisées par la suite.

PROPOSITION 1. — Un processus stationnaire $\{X_n\}$ est à rang complet, si et seulement si

$$(2.9) \quad \int_{-\pi}^{\pi} |\log \Delta(F'_{xx}(\lambda))| d\lambda < \infty.$$

PROPOSITION 2. — Pour un processus à rang complet, le déterminant G de l'innovation vérifie la relation

$$(2.10) \quad \Delta(G) = \exp \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \Delta(F'_{xx}(\lambda)) d\lambda \right]$$

2.3. CALCUL DU DÉBIT D'ENTROPIE D'UN PROCESSUS LAPLACIEN STATIONNAIRE.

Soit une suite quelconque $X_1, X_2, \dots, X_n$ de n vecteurs laplaciens à q dimensions. Désignons par $\Gamma_{i,j}$ ($1 \leq i, j \leq n$) la matrice de covariance de X_i et X_j et supposons que le déterminant de la matrice

$$(3.1) \quad \Gamma^n = \begin{vmatrix} \Gamma_{11} & \Gamma_{12} & \dots & \Gamma_{1n} \\ \Gamma_{21} & \Gamma_{22} & \dots & \Gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{n1} & \Gamma_{n2} & \dots & \Gamma_{nn} \end{vmatrix}, \quad \Gamma_{ij} = \Gamma_{ji}^*$$

soit différent de zéro. La projection $\mathbf{X}_n$ de X_n sur le sous-espace engendré par $X_1, X_2, \dots, X_{n-1}$ peut se mettre sous la forme

$$(3.2) \quad \mathbf{X}_n = \sum_{j=1}^{n-1} A_j X_j.$$

Si l'on exprime la relation d'orthogonalité

$$(3.3) \quad (X_n - \mathbf{X}_n, X_k) = 0 \quad (1 \leq k \leq n-1)$$

nous obtenons un système de $n-1$ équations linéaires à $n-1$ inconnues qui sont les A_j ; c'est-à-dire

$$(3.4) \quad \Gamma_{n,k} - \sum_{j=1}^{n-1} A_j \Gamma_{j,k} = 0 \quad (1 \leq k \leq n-1).$$

Nous nous intéressons ici à la valeur de $(X_n - \mathbf{X}_n, X_n - \mathbf{X}_n)$ non pas aux inconnues A_j . Nous avons, compte tenu de la relation d'orthogonalité (3.5),

$$(3.5) \quad (X_n - \mathbf{X}_n, X_n - \mathbf{X}_n) = \Gamma_{n,n} - \sum_{j=1}^{n-1} A_j \Gamma_{j,n}.$$

Posons $A = (A_1 \ A_2 \dots A_{n-1})$; $\gamma = (\Gamma_{n,1} \ \Gamma_{n,2} \dots \ \Gamma_{n,n-1})$

Les équations (3.4) se traduisent par l'équation matricielle

$$A \Gamma^{n-1} = \gamma, \quad \text{d'où} \quad A = \gamma \cdot (\Gamma^{n-1})^{-1}.$$

En reportant A dans (3.5), nous trouvons

$$(3.6) \quad (X_n - \mathbf{X}_n, X_n - \mathbf{X}_n) = \Gamma_{n,n} - \gamma (\Gamma^{n-1})^{-1} \gamma^*.$$

A partir de cette dernière expression du gramien de l'erreur de prévision $X_n - \mathbf{X}_n$, nous pouvons calculer le déterminant de ce gramien de la manière suivante. Nous avons

$$\Gamma^n = \begin{vmatrix} \Gamma^{n-1} & \gamma^* \\ \gamma & \Gamma_{n,n} \end{vmatrix} \quad \text{et soit la matrice } C = \begin{vmatrix} (\Gamma^{n-1})^{-1} & 0 \\ 0 & (\Gamma_{n,n})^{-1} \end{vmatrix}. \quad \text{Faisons le produit}$$

$$\Gamma^n \cdot C = \begin{vmatrix} I & \gamma^* (\Gamma_{n,n})^{-1} \\ \gamma (\Gamma^{n-1})^{-1} & I \end{vmatrix}$$

où I est la matrice unité. Il suit de l'expression de cette dernière matrice

$$\Delta (\Gamma^n C) = \frac{\Delta (\Gamma^n)}{\Delta (\Gamma^{n-1}) \Delta (\Gamma_{n,n})} = \Delta (I - \gamma (\Gamma^{n-1})^{-1} \gamma^* (\Gamma_{n,n})^{-1}). \quad \text{D'où}$$

$$(3.7) \quad \Delta (X_n - \mathbf{X}_n, X_n - \mathbf{X}_n) = \Delta (\Gamma_{n,n} - \gamma (\Gamma^{n-1})^{-1} \gamma^*) = \frac{\Delta (\Gamma^n)}{\Delta (\Gamma^{n-1})}.$$

Soit X un vecteur-colonne réel, laplacien, à q dimensions. Ses composantes $x^1, x^2, \dots, x^q$ possèdent la densité de probabilité jointe

$$p(x^1, \dots, x^q) = [(2\pi)^q \Delta(\Gamma)]^{-1/2} \exp \left\{ -\frac{1}{2} \tilde{X}(\Gamma)^{-1} X \right\},$$

où Γ est la matrice de covariance et $\tilde{X}$ le transposé de X . L'entropie différentielle de X est donnée par

$$H(X) = \frac{1}{2} \log [(2\pi e)^q \Delta(\Gamma)].$$

Par conséquent, d'après l'expression (3.7), si les vecteurs $X_1, X_2, \dots, X_n$ sont laplaciens, $X_n - X_n$ aura pour entropie différentielle

$$(3.8) \quad H(X_n - X_n) = \frac{1}{2} \log(2\pi e)^q \frac{\Delta(\Gamma^n)}{\Delta(\Gamma^{n-1})}.$$

Si maintenant, nous cherchions à calculer l'entropie différentielle de X_n , connaissant $X_{n-1}, \dots, X_1$, en prenant la moyenne du logarithme de sa densité de probabilité conditionnelle

$$\left[(2\pi)^q \frac{\Delta(\Gamma^n)}{\Delta(\Gamma^{n-1})} \right]^{-1/2} \cdot \exp \left\{ -\frac{1}{2} \tilde{X}^n (\Gamma^n)^{-1} X^n + \frac{1}{2} \tilde{X}^{n-1} (\Gamma^{n-1})^{-1} X^{n-1} \right\} \quad \text{où } X^n = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

les X_j ($1 \leq j \leq n$) étant considérés comme des vecteurs-colonne, nous trouverions

$$(3.9) \quad H(X_n/X_{n-1}, \dots, X_1) = \frac{1}{2} \log(2\pi e)^q \frac{\Delta(\Gamma^n)}{\Delta(\Gamma^{n-1})}.$$

La comparaison de (3.8) et de (3.9) nous amène à la conclusion suivante :

THÉORÈME B. — Soient $X_1, \dots, X_n$ des vecteurs aléatoires laplaciens à q dimensions, tous corrélés entre eux, et supposons que $\Delta(\Gamma^n) \neq 0$. L'entropie conditionnelle de l'un d'entre eux, connaissant les autres est égale à l'entropie du vecteur obtenu en prenant la différence entre ce vecteur et sa projection sur le sous-espace linéaire engendré par les autres. Le résultat est valable si les vecteurs considérés sont complexes.

Ce théorème est évidemment une conséquence directe des propriétés des mesures laplaciennes, l'orthogonalité se confondant avec l'indépendance. Il entraîne alors que, si une suite stationnaire laplacienne est à rang complet, l'entropie conditionnelle

$$\lim_{p \rightarrow \infty} H(X_n/X_{n-1}, \dots, X_{n-p})$$

existe et elle est indépendante de n et égale à l'entropie de l'innovation du processus. Cette remarque, le théorème A et l'expression (2.10) entraînent le résultat suivant :

THÉORÈME C. — Soit X_n un processus aléatoire stationnaire, laplacien, à q dimensions. Si le processus est à rang complet, son débit d'entropie existe et il est donné par

$$(3.10) \quad \bar{H}(X) = \frac{q}{2} \log(2\pi e) + \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \Delta \{ F'_{xx}(\lambda) \} d\lambda.$$

Le résultat obtenu ici, nous amène tout naturellement à la remarque suivante :

REMARQUE. — Considérons la matrice Γ^n concernant $n+1$ éléments successifs de la suite X_n

$$\Gamma^n = \begin{pmatrix} \Gamma_0 & \Gamma_1 & \Gamma_2 & \dots & \Gamma_n \\ \Gamma_1^* & \Gamma_0 & \Gamma_1 & \dots & \Gamma_{n-1} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \Gamma_n^* & \Gamma_{n-1}^* & \Gamma_{n-2}^* & \dots & \Gamma_0 \end{pmatrix} \quad \text{où } \Gamma_j \text{ est donné par } \Gamma_j = \int_{-\pi}^{\pi} e^{ij\lambda} dF_{xx}(\lambda)$$

$$\begin{aligned} \text{La relation limite : } & \lim_{n \rightarrow \infty} \frac{1}{n+1} H(X_0, X_{-1}, \dots, X_n) = \lim_{n \rightarrow \infty} H(X_0/X_{-1}, \dots, X_{-n}) \\ & = \frac{q}{2} \log(2\pi e) + \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \Delta \{ F'_{xx}(\lambda) \} d\lambda \end{aligned}$$

généralise un résultat dû à Szego (cf. Grenander et Szego [1]), à savoir

$$\lim_{n \rightarrow \infty} [\Delta(\Gamma)]^{\frac{1}{n+1}} = \lim_{n \rightarrow \infty} \frac{\Delta(\Gamma^n)}{\Delta(\Gamma^{n-1})} = \exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \Delta \{ F'_{xx}(\lambda) \} d\lambda$$

Ce résultat était connu dans le cas où les éléments Γ_j sont des nombres complexes, non des matrices comme ici.

2.4. FILTRAGE LINÉAIRE DES PROCESSUS STATIONNAIRES ET DÉBIT D'INFORMATION TRANSMISE ENTRE DEUX PROCESSUS LAPLACIENS, STATIONNAIRES ET STATIONNAIREMENT CORRÉLÉS.

Le débit d'information entre deux processus $\{X_n\}$ et $\{Y_n\}$ sera calculé par la formule

$$\bar{I}(X, Y) = \bar{H}(X) - \bar{H}(X/Y).$$

Pour calculer le débit d'entropie conditionnelle du processus $\{X_n\}$ connaissant le processus $\{Y_n\}$ nous tiendrons compte des possibilités pratiques du temps disponible à la détection du signal reçu. Il est très souvent impossible d'attendre de recevoir toute la suite $\{Y_n\}$ pour commencer à estimer la valeur des signaux émis. On peut tolérer seulement un retard de temps p (ici, p sera un entier). Voici comment le problème peut se poser dans le cas où $\{X_n\}$ et $\{Y_n\}$ sont deux processus vectoriels laplaciens, stationnaires et stationnairement corrélés. Ayant connu les valeurs des variables $Y_n, Y_{n-1}, \dots$ on cherche la meilleure estimation linéaire de X_{n-p} . Nous supposons ici, que les processus $\{X_n\}$ et $\{Y_n\}$ sont tous les deux à q dimensions (uniquement pour simplifier l'écriture; car, la démonstration des résultats n'en dépend pas). Aux processus $\{X_n\}$ et $\{Y_n\}$ correspondent respectivement deux fonctions $X(\lambda)$ et $Y(\lambda)$ à accroissements non corrélés telles que

$$X_n = \int_{-\pi}^{\pi} e^{in\lambda} dX(\lambda); \quad Y_n = \int_{-\pi}^{\pi} e^{in\lambda} dY(\lambda) \\ (dX(\lambda), dY(\mu)) = 0 \quad \text{pour } \lambda \neq \mu$$

Nous supposons que leurs fonctions de distribution sont absolument continues, c'est-à-dire que

$$(dX(\lambda), dX(\lambda)) = f_{xx}(\lambda) d\lambda, \quad (dX(\lambda), dY(\lambda)) = f_{xy}(\lambda) d\lambda,$$

et que

$$(dX(\lambda), dY(\lambda)) = f_{xy}(\lambda) d\lambda = f_{yx}^*(\lambda) d\lambda.$$

Nous supposons, de plus, que $\{Y_n\}$ est à rang complet.

Soit $\hat{X}_{n-p}$ la projection de X_{n-p} sur le sous-espace linéaire engendré par les éléments Y_j ($j \leq n$) du processus $\{Y_n\}$. $\hat{X}_{n-p}$ peut se mettre sous la forme

$$(4.1) \quad \hat{X}_{n-p} = \int_{-\pi}^{\pi} e^{in\lambda} \varphi_{X/Y}(\lambda) dY(\lambda)$$

où la fonction matricielle $\varphi_{X/Y}(\lambda)$ n'a pas d'exposants positifs dans son développement de Fourier, c'est-à-dire qu'elle est de la forme

$$\varphi_{X/Y}(\lambda) = \sum_{j=0}^{\infty} A_j e^{-ij\lambda}$$

et elle vérifie la relation

$$\int_{-\pi}^{\pi} \text{Tr} (\varphi_{X/Y}(\lambda) f_{YY}(\lambda) \varphi_{X/Y}^*(\lambda)) d\lambda < \infty$$

D'après le théorème de projection, nous avons $X_{n-p} - \hat{X}_{n-p} \perp Y_j$, quel que soit $j \leq n$, et, en particulier,

$$(4.2) \quad X_{n-p} - \hat{X}_{n-p} \perp \hat{X}_{n-p}.$$

Cette dernière entraîne

$$(4.3) \quad \int_{-\pi}^{\pi} [e^{-ip\lambda} f_{XY}(\lambda) - \varphi_{X/Y}(\lambda) f_{YY}(\lambda)] \varphi_{X/Y}^*(\lambda) d\lambda = 0.$$

Soit $\Phi_{YY}(\lambda)$ la fonction génératrice du processus $\{Y_n\}$. $\Phi_{YY}(\lambda)$ n'a pas d'exposants positifs dans son développement de Fourier, et

$$(4.4) \quad f_{YY}(\lambda) = \Phi_{YY}(\lambda) \Phi_{YY}^*(\lambda).$$

La condition (4.3) peut aussi se mettre sous la forme :

$$(4.5) \quad \int_{-\pi}^{\pi} \{ e^{-ip\lambda} f_{x|y}(\lambda) [\Phi_{yy}^*(\lambda)]^{-1} - \varphi_{x|y}(\lambda) \Phi_{yy}(\lambda) \} \times \Phi_{yy}^*(\lambda) \varphi_{x|y}^*(\lambda) d\lambda = 0$$

Cette condition sera vérifiée, si l'expression entre accolades n'a que des exposants positifs dans son développement de Fourier. Considérons la série de Fourier

$$(4.6) \quad f_{xy}(\lambda) [\Phi_{yy}^*(\lambda)]^{-1} = \sum_{-\infty}^{\infty} B_m e^{-im\lambda}$$

qui converge en norme euclidienne, en ce sens que

$$\lim_{M \rightarrow \infty} \int_{-\pi}^{\pi} \text{Tr} (f_{xy}(\lambda) [\Phi_{yy}^*(\lambda)]^{-1} [\Phi_{yy}(\lambda)]^{-1} f_{yx}(\lambda) - \sum_{-M}^M B_m e^{-im\lambda}) d\lambda = 0.$$

La relation (4.5) sera vérifiée, si et seulement si

$$(4.7) \quad \varphi_{x|y}(\lambda) \cdot \Phi_{yy}(\lambda) = \sum_{m=-p}^{\infty} B_m e^{-i(m+p)\lambda}; \quad \text{d'où}$$

$$(4.8) \quad \varphi_{x|y}(\lambda) = \left[\sum_{m=-p}^{\infty} B_m e^{-i(m+p)\lambda} \right] [\Phi_{yy}(\lambda)]^{-1}$$

Le processus d'erreur $Z_{n-p} = X_{n-p} - \hat{X}_{n-p}$ est aussi stationnaire pour p constant, avec la densité de distribution spectrale

$$(4.9) \quad f_{zz}(\lambda) = f_{xx}(\lambda) - \varphi_{x|y}(\lambda) f_{yy}(\lambda) \varphi_{x|y}^*(\lambda),$$

qui est aussi une matrice hermitienne non négative. Ou bien, si l'on pose

$$(4.10) \quad B_p(\lambda) = \sum_{m=-p}^{\infty} B_m e^{-i(m+p)\lambda} \quad \text{nous avons}$$

$$(4.11) \quad f_{zz}(\lambda) = f_{xx}(\lambda) - B_p(\lambda) B_p^*(\lambda).$$

Pour le calcul du débit d'information, nous utiliserons le fait que, le récepteur étant obligé de restituer les signaux avec un retard fini p , il lui est impossible d'utiliser la liaison stochastique qui peut effectivement exister entre X_{n-p} et les signaux reçus après l'instant n . Tout se passe comme si la suite $\{Z_{n-p}\}$ calculée plus haut pour p constant était indépendante de la suite $\{Y_n\}$ de signaux reçus. Alors, d'après le théorème B, le débit d'entropie conditionnelle du processus $\{X_n\}$ calculé jusqu'à l'instant $n - p$, connaissant $\{Y_n\}$ jusqu'à l'instant n , sera égal au débit d'entropie du processus d'erreur $\{Z_n\}$. Nous avons donc :

$$(4.12) \quad \bar{H}(X/Y) = \frac{q}{2} \log(2\pi e) + \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \Delta [f_{xx}(\lambda) - B_p(\lambda) B_p^*(\lambda)] d\lambda.$$

D'où l'expression du débit d'information entre les deux processus.

THÉOREME D. — Lorsque le retard admissible à la réception est égal à p , le débit d'information entre deux processus vectoriels laplaciens, stationnaires, stationnairement corrélés entre eux et tous deux à rang complet, est donné par la formule

$$\bar{I}_p(X, Y) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \frac{\Delta [f_{xx}(\lambda)]}{\Delta [f_{xx}(\lambda) - B_p(\lambda) B_p^*(\lambda)]} d\lambda$$

où B_p est donné par (4.6) et (4.10).

REMARQUE 1. — Si l'on fait tendre p vers l'infini, nous retrouvons un résultat déjà connu (cf. Pinsker [2], Korezlioglu [3]). Nous avons alors

$$(4.14) \quad \lim_{p \rightarrow \infty} \bar{I}_p(X, Y) = \bar{I}(X, Y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \frac{\Delta [f_{xx}(\lambda)]}{\Delta [f_{xx}(\lambda) - f_{xy}(\lambda) [f_{yy}(\lambda)]^{-1} f_{yx}(\lambda)]} d\lambda$$

Il est aisé de voir que cette expression est symétrique par rapport aux deux processus $\{X_n\}$ et $\{Y_n\}$; ce qui n'était pas le cas de l'expression (4.13). Les relations suivantes donnent des expressions

équivalentes pour $\bar{I}(X, Y)$.

$$\begin{aligned} \frac{\Delta \left| \frac{f_{xx}(\lambda) f_{xy}(\lambda)}{f_{yx}(\lambda) f_{yy}(\lambda)} \right|}{\Delta(f_{xx}(\lambda)) \Delta(f_{yy}(\lambda))} &= \Delta(I - f_{xy}(\lambda) [f_{yy}(\lambda)]^{-1} f_{yx}(\lambda) [f_{xx}(\lambda)]^{-1}) \\ &= \Delta(I - f_{yx}(\lambda) [f_{xx}(\lambda)]^{-1} f_{xy}(\lambda) [f_{yy}(\lambda)]^{-1}) \\ &= \frac{\Delta(f_{xx}(\lambda))}{\Delta(f_{xx}(\lambda) - f_{xy}(\lambda) [f_{yy}(\lambda)]^{-1} f_{yx}(\lambda))} = \frac{\Delta(f_{yy}(\lambda))}{\Delta(f_{yy}(\lambda) - f_{yx}(\lambda) [f_{xx}(\lambda)]^{-1} f_{xy}(\lambda))} \end{aligned}$$

REMARQUE 2. — En fait, le débit d'information que l'on pourrait transmettre entre $\{X_n\}$ pour $n \leq -p$ et $\{Y_n\}$ pour $n \leq 0$ est encore égal à (4.14), d'après la remarque faite à la fin du § 1. Mais alors, il n'y a plus de restriction au retard. En effet, en remplaçant $\{X_n\}$ par le processus

$$X'_n = \int_{-\pi}^{\pi} e^{in\lambda} e^{-ip\lambda} dX(\lambda)$$

de même densité spectrale que $\{X_n\}$, on voit que l'expression (4.14) reste inchangée.

REMARQUE 3. — Si les deux processus sont respectivement à une dimension, la formule (4.14) se réduit à la formule déjà donnée par Pinsker [1].

$$(4.15) \quad \bar{I}(x, y) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \frac{f_{xx}(\lambda) f_{yy}(\lambda)}{f_{xx}(\lambda) f_{yy}(\lambda) - |f_{xy}(\lambda)|^2} d\lambda.$$

Si les deux processus sont à paramètre continu, on peut, en utilisant le théorème d'échantillonnage de Shannon (cf. Pinsker [1], Guelfand et Jaglom [1]) calculer le débit d'information entre les deux processus. On trouve une formule analogue à (4.15)

$$(4.16) \quad \bar{I}(x, y) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \log \frac{f_{xx}(\lambda) f_{yy}(\lambda)}{f_{xx}(\lambda) f_{yy}(\lambda) - |f_{xy}(\lambda)|^2} d\lambda$$

REMARQUE 4. — Avec les hypothèses faites pour l'établissement de l'expression (4.11) la condition nécessaire et suffisante pour que la suite $\{X_n\}$ ($n \leq -p$) soit subordonnée à la suite $\{Y_n\}$ ($n \leq 0$) est que

$$f_{xx}(\lambda) - B_p(\lambda) B_p^*(\lambda) = 0.$$

En fait, pour $p = \infty$, cette condition peut s'exprimer indépendamment des hypothèses faites sur la continuité absolue des fonctions de distribution spectrales. Nous avons en effet

$$\varphi_{x|y}(\lambda) = dF_{xy}(\lambda) [dF_{yy}(\lambda)]^{-1}, \quad \text{d'où} \quad dF_{zz}(\lambda) = dF_{xx}(\lambda) - dF_{xy}(\lambda) [dF_{yy}(\lambda)]^{-1} dF_{yx}(\lambda).$$

La condition de subordination s'exprime alors, par

$$\int_{-\pi}^{\lambda} dF_{xx}(\mu) - dF_{xy}(\mu) [dF_{yy}(\mu)]^{-1} dF_{yx}(\mu) = 0.$$

(cf. Karhunen [2], Kolmogorov, pour le cas des processus à une dimension).

2.5. APPLICATIONS.

2.5.1. Un cas particulier concernant les processus laplaciens non stationnaires.

Nous considérons deux processus laplaciens $\{x_n\}$ et $\{y_n\}$ non stationnaires, tous deux indépendants de leurs passés infiniment lointains et dont l'un au moins est purement non déterministe, c'est-à-dire que ses instants d'indétermination coïncident avec l'ensemble des entiers. Nous avons vu, dans le premier chapitre, que $\{x_n\}$ et $\{y_n\}$ admettent les représentations intégrales suivantes :

$$(5.1) \quad \begin{aligned} x_n &= \int_{-\pi}^{\pi} e^{in\lambda} \gamma_n^x(\lambda) d\xi(\lambda) & \|d\xi(\lambda)\|^2 &= \frac{d\lambda}{2\pi} \\ y_n &= \int_{-\pi}^{\pi} e^{in\lambda} \gamma_n^y(\lambda) d\eta(\lambda) & \|d\eta(\lambda)\|^2 &= \frac{d\lambda}{2\pi} \end{aligned}$$

Supposons que la condition

$$E[d\xi(\lambda) \cdot d\bar{\eta}(\mu)] = 0 \quad \text{pour } \lambda \neq \mu$$

soit vérifiée et posons

$$E[d\xi(\lambda) \cdot d\bar{\eta}(\lambda)] = \frac{1}{2\pi} f_{\xi\eta}(\lambda) d\lambda.$$

Nous nous proposons de calculer le débit d'information transmise entre les deux processus, mais en admettant un retard infini à la réception. Nous calculons, tout d'abord, le débit d'entropie du processus $\{x_n\}$. L'entropie de x_n , connaissant le passé du processus est donnée par

$$(5.2) \quad H(x_n/x_{n-1}, x_{n-2}, \dots) = \frac{1}{2} \log 2\pi e (c_{n,n}^x)^2$$

où $(c_{n,n}^x)^2$ est la variance de l'innovation du processus $\{x_n\}$ à l'instant n . Nous pouvons vérifier facilement la relation suivante :

$$(5.3) \quad \bar{H}(x) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} H(x_N, \dots, x_{-N}) = \frac{1}{2} \log 2\pi e + \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N \log c_{n,n}^x.$$

D'autre part, pour le calcul du débit d'entropie conditionnelle du processus $\{x_n\}$ connaissant $\{y_n\}$, il nous faut calculer la projection de x_n sur $\mathcal{E}(y)$. Désignons par $\mathbf{x}_n$ cette projection. Nous l'écrivons sous la forme intégrale :

$$\mathbf{x}_n = \int_{-\pi}^{\pi} e^{in\lambda} G_n(\lambda) d\eta(\lambda).$$

La relation d'orthogonalité $E[(x_n - \mathbf{x}_n) \cdot \bar{\mathbf{x}}_n] = 0$ entraîne $G_n(\lambda) = \gamma_n^x(\lambda) f_{\xi\eta}(\lambda)$.

La variance de $x_n - \mathbf{x}_n = z_n$ est alors donnée par

$$E|x_n - \mathbf{x}_n|^2 = \int_{-\pi}^{\pi} |\gamma_n^x(\lambda)|^2 (1 - |f_{\xi\eta}(\lambda)|^2) \frac{d\lambda}{2\pi}.$$

Tout se passe comme si le processus d'erreur $\{z_n\} = \{x_n - \mathbf{x}_n\}$ était subordonné à un processus stationnaire de densité spectrale $\frac{1}{2\pi} (1 - |f_{\xi\eta}(\lambda)|^2)$ et nous savons que la variance c_0^z de l'innovation de ce processus est donnée par

$$c_0^z = \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log (1 - |f_{\xi\eta}(\lambda)|^2) d\lambda \right\}.$$

Comme la suite de fonctions $\gamma_n^x(\lambda)$ forme un système complet dans $L_2(-\pi, \pi)$, la variance de l'innovation à l'instant n du processus d'erreur $\{z_n\}$ est égale à

$$(c_{n,n}^x)^2 \cdot \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log (1 - |f_{\xi\eta}(\lambda)|^2) d\lambda \right\}.$$

Le débit d'entropie conditionnelle de $\{x_n\}$, connaissant $\{y_n\}$ étant égal au débit d'entropie du processus d'erreur $\{z_n\}$, nous avons :

$$(5.5) \quad \begin{aligned} \bar{H}(x/y) &= \lim_{N \rightarrow \infty} \frac{1}{2N+1} H(z_N, \dots, z_{-N}) \\ &= \frac{1}{2} \log 2\pi e + \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N \log c_{n,n}^x + \frac{1}{4\pi} \int_{-\pi}^{\pi} \log (1 - |f_{\xi\eta}(\lambda)|^2) d\lambda \\ &= \bar{H}(x) + \frac{1}{4\pi} \int_{-\pi}^{\pi} \log (1 - |f_{\xi\eta}(\lambda)|^2) d\lambda. \end{aligned} \quad \text{D'où}$$

$$(5.6) \quad \bar{I}(x, y) = \bar{H}(x) - \bar{H}(x/y) = -\frac{1}{4\pi} \int_{-\pi}^{\pi} \log (1 - |f_{\xi\eta}(\lambda)|^2) d\lambda.$$

Dans le cas où les processus $\{x_n\}$ et $\{y_n\}$ sont stationnaires, les expressions (5.1) peuvent s'écrire

$$(5.7) \quad x_n = \int_{-\pi}^{\pi} e^{in\lambda} \gamma^x(\lambda) d\xi(\lambda); \quad y_n = \int_{-\pi}^{\pi} e^{in\lambda} \gamma^y(\lambda) d\eta(\lambda)$$

En reportant dans la formule (4.15) les densités spectrales

$$f_{xx}(\lambda) = \frac{1}{2\pi} |\gamma^x(\lambda)|^2, f_{yy}(\lambda) = \frac{1}{2\pi} |\gamma^y(\lambda)|^2 \text{ et } f_{xy}(\lambda) = \frac{1}{2\pi} \gamma^x(\lambda) \bar{\gamma}^y(\lambda) f_{\xi\eta}(\lambda)$$

qui se déduisent des expressions (5.7), nous voyons que la formule (4.15) se réduit à la formule (5.6).

2.5.2. Confirmation de la formule de capacité de Shannon.

Les formules (4.15) et (4.16) entraînent que, si les densités spectrales $\{x_n\}$ et $\{y_n\}$ sont nulles en dehors d'un intervalle de fréquence $[-w, +w]$, le débit d'information entre les deux processus

peut s'exprimer par la formule

$$(5.8) \quad \bar{I}(x, y) = \frac{1}{2} \int_{-w}^w \log \frac{f_x(v) f_y(v)}{f_x(v) f_y(v) - |f_{xy}(v)|^2} dv$$

où les densités spectrales sont exprimées par rapport à la variable fréquence $v = \frac{\lambda}{2\pi}$.

Supposons que le signal reçu $\{y_n\}$ soit la somme du signal émis $\{x_n\}$ et d'un bruit laplacien $\{b_n\}$ indépendant de $\{x_n\}$ et de densité spectrale $f_b(v)$. La formule (5.8) devient :

$$(5.9) \quad \bar{I}(x, y) = \int_0^w \log \left(1 + \frac{f_x(v)}{f_b(v)} \right) dv.$$

Si, en plus, les processus $\{x_n\}$ et $\{b_n\}$ ont des densités spectrales constantes, respectivement $S/2w$ et $B/2w$ dans la bande $[-w, +w]$, le débit d'information transmise sera

$$(5.10) \quad \bar{I} = w \log \left(1 + \frac{S}{B} \right).$$

Shannon a montré que, si un canal stationnaire de bande passante $[-w, +w]$ est perturbé par un bruit blanc de puissance moyenne B et si la puissance moyenne de la source est égale à S , le maximum du débit d'information, que l'on peut transmettre par le canal ainsi défini, est obtenu lorsque la source aussi est un bruit blanc de puissance moyenne S . La formule (5.10) donne alors la valeur de ce maximum. C'est, par définition, la capacité de ce canal, que l'on peut retrouver de la manière suivante.

Considérons d'abord, la formule

$$\bar{I}(x, y) = \bar{H}(y) - \bar{H}(y/x).$$

On a $\bar{H}(y/x) = \bar{H}(b)$, car le bruit est indépendant de $\{x_n\}$. On peut démontrer que le débit d'entropie de la source stationnaire $\{y_n\}$ de densité spectrale $f_{yy}(v) = f_{xx}(v) + B/2w$ est maximum lorsque la source $\{x_n\}$ est laplacienne. On trouve alors :

$$\bar{I}(x, y) = \int_0^w \log \left(1 + \frac{f_x(v)}{B/2w} \right) dv.$$

Le maximum de cette expression, compte tenu de la condition $\int_{-w}^w f_x(v) dv = S$ est obtenu par

$$f_x(v) = S/2w.$$

Remarque. — Cette expression de la capacité $C = w \log \left(1 + \frac{S}{B} \right)$ est uniquement basée sur la définition du débit d'information que nous avons donnée conformément à la théorie de Shannon. Cette notion de maximum du débit d'information que l'on peut transmettre à travers un canal est susceptible d'une interprétation plus concrète, grâce à une version généralisée du *lemme fondamental de Feinstein* donnée par A. Perez.

Rappelons tout d'abord, la notion de ε -discernabilité à travers un canal.

Soient $(X, \mathfrak{X})$ et $(Y, \mathfrak{Y})$ deux espaces mesurables. Soit $(\nu_x, x \in X)$ une famille de mesures de probabilité sur $\mathfrak{Y}$ telle que, pour tout ensemble $F \in \mathfrak{Y}$, $\nu_x(F)$ est une fonction de x , mesurable par rapport à la σ -algèbre $\mathfrak{X}$. On dit qu'un tel système constitue un canal avec la σ -algèbre d'entrée $\mathfrak{X}$ et la σ -algèbre de sortie $\mathfrak{Y}$ et on le désigne par $(\mathfrak{X}, \nu_x, \mathfrak{Y})$.

Soit $\varepsilon > 0$ arbitrairement petit. L'ensemble de points $\{x_\gamma : x_\gamma \in X, \gamma \in \Gamma\}$ (où Γ est un ensemble numérique) est appelé par Perez, un groupe ε -discernable à travers un canal $(\mathfrak{X}, \nu_x, \mathfrak{Y})$, s'il est possible de faire correspondre à chaque point x_γ un ensemble non vide $F_\gamma \in \mathfrak{Y}$, $\gamma \in \Gamma$, de façon que les ensembles F_γ soient disjoints et que

$$\nu_{x_\gamma}(F_\gamma) > 1 - \varepsilon.$$

Le supremum des puissances de tous les groupes ε -discernables à travers un canal est appelé la puissance de ε -discernabilité du canal.

Considérons un canal stationnaire $(\prod_{i=-\infty}^{\infty} \mathfrak{X}_i, \nu_x, \prod_{i=-\infty}^{\infty} \mathfrak{Y}_i)$, où $\mathfrak{X}_i = \mathfrak{X}_0$ et $\mathfrak{Y}_i = \mathfrak{Y}_0$ quel que soit i , et désignons par C sa capacité. La version généralisée du lemme fondamental de Feinstein s'exprime de la manière suivante :

Pour tout $0 < \varepsilon < \frac{1}{2}$ arbitrairement petit, on peut trouver un entier $n_0(\varepsilon)$ tel que pour $n \geq n_0(\varepsilon)$,

la puissance $N_n(\varepsilon)$ de ε -discernabilité du canal $[(\mathfrak{H}_0)^n, \nu_x, (\mathfrak{H}_0)^n]$ vérifie l'inégalité

$$N_n(\varepsilon) > e^{n(C-2\varepsilon)}.$$

Dans le cas d'un canal de bande de fréquence limitée à $[-w, w]$ et perturbé par un bruit blanc de puissance moyenne B et indépendant des signaux émis, lorsque la puissance de ces derniers est égale à S , le choix des ensembles ε -discernables se ramène toujours à celui d'une distance optimale dans l'espace $L_2(-w, +w)$ parmi les fonctions ayant une norme donnée comme l'a montré Bétoux.

Bétoux a donné une expression approchée de la capacité d'un tel canal, à partir de la valeur de la puissance de ε -discernabilité des sous-espaces linéaires à dimensions finies, dans le cas où $S/B \ll 1$. Ce qui fait par conséquent

$$C \simeq w S/B.$$

2.5.3. Transmission des innovations.

L'information transmise à travers un canal de largeur de bande limitée et perturbée par un bruit blanc indépendant du signal, atteint sa valeur maximale, lorsque le signal émis est aussi un bruit blanc. Par conséquent, pour l'utilisation du canal avec le maximum de rendement, la source doit émettre des grandeurs laplaciennes indépendantes. La transmission de la suite orthonormale $\{\xi_n\}$, définie à partir des innovations d'un processus laplacien sera la meilleure utilisation du canal, si l'estimation de ce dernier processus peut se faire avec la précision voulue. Considérons, par exemple, un processus stationnaire purement non déterministe défini par le développement canonique

$$x_n = \sum_{m=0}^{\infty} c_m \xi_{n-m}.$$

Soient S la puissance moyenne de la source et B celle du bruit dans $[-\pi, \pi]$, la source émettant la suite $\{\sqrt{S} \xi_n\}$. Le signal reçu sera

$$\{y_n\} = \{\sqrt{S} \xi_n + b_n\}.$$

Soit $\sqrt{S} \hat{\xi}_n$ l'estimation linéaire de $\sqrt{S} \xi_n$ connaissant $\{y_n\}$, nous posons

$$\sqrt{S} \hat{\xi}_n = \int_{-\pi}^{\pi} e^{in\lambda} \gamma(\lambda) dY(\lambda); \quad \xi_n = \int_{-\pi}^{\pi} e^{in\lambda} d\xi(\lambda); \quad y_n = \int_{-\pi}^{\pi} e^{in\lambda} dY(\lambda).$$

La relation d'orthogonalité $(\sqrt{S} \xi_n - \sqrt{S} \hat{\xi}_n) \perp \sqrt{S} \hat{\xi}_n$ entraîne

$$\gamma(\lambda) = \frac{S}{B + S}, \quad \text{d'où}$$

$$(5.11) \quad \|\xi_n - \hat{\xi}_n\|^2 = \frac{B}{S + B}.$$

Comme

$$\|x_n - \mathbf{x}_n\|^2 = \sum_{m=0}^{\infty} |c_m|^2 \|\hat{\xi}_{n-m} - \xi_{n-m}\|^2,$$

où $\mathbf{x}_n$ est l'estimation de x_n connaissant $\{y_n\} = \{\sqrt{S} \xi_n + b_n\}$, nous avons

$$\frac{\|x\|^2}{\|x - \mathbf{x}\|^2} = 1 + \frac{S}{B}$$

Si, par contre, le signal émis était directement le processus $\{\mathbf{x}_n\}$, avec une puissance moyenne S et une densité spectrale $f_x(\lambda)$ telle que

$$(5.12) \quad \int_{-\pi}^{\pi} f_x(\lambda) d\lambda = S,$$

nous aurions

$$(5.13) \quad \|x_n - \mathbf{x}_n\|^2 = \int_{-\pi}^{\pi} \frac{f_x(\lambda) \cdot B/2\pi}{f_x(\lambda) + B/2\pi} d\lambda.$$

La quantité (5.11) est donc à comparer à

$$(5.14) \quad \frac{\|x_n\|^2}{\|x_n - \mathbf{x}_n\|^2} = \frac{S}{B} \cdot \frac{1}{\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f_x(\lambda) d\lambda}{f_x(\lambda) + B/2\pi}}.$$

Cherchons maintenant un extremum de (5.13) tenant compte de (5.12). Posons

$$J(f) = \int_{-\pi}^{\pi} \frac{f_x(\lambda)}{f_x(\lambda) + B/2\pi} d\lambda - \mu \int_{-\pi}^{\pi} f_x(\lambda) d\lambda.$$

Pour que $\delta J(f) = \int_{-\pi}^{\pi} \frac{B}{2\pi} \cdot \frac{\delta f_x(\lambda)}{[f_x(\lambda) + B/2\pi]^2} d\lambda - \mu \int_{-\pi}^{\pi} \delta f_x(\lambda) d\lambda = 0,$

nous avons $f_x(\lambda) = \sqrt{\frac{B}{2\pi\mu}} - \frac{B}{2\pi}.$

Comme $\delta^2 J(f) < 0$, $J(\sqrt{\frac{B}{2\pi\mu}} - \frac{B}{2\pi})$ est un maximum de $J(f)$.

Si l'on tient compte de la condition (5.12), on trouve

$$\mu = 2\pi \cdot B / (S + B)^2. \quad \text{D'où}$$

$$\min_{f_x} S/B \cdot \frac{1}{\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f_x(\lambda) d\lambda}{f_x(\lambda) + B/2\pi}} = 1 + S/B$$

La précision relative de l'estimation de $\{x_n\}$ est donc meilleure que lorsque seules ses innovations sont transmises. Ce résultat pourrait se traduire en langage des télécommunications, en disant que la diminution de la redondance à l'émission diminue la précision relative de l'estimation à la réception. Nous voyons donc que la transmission des innovations ne peut convenir que si le bruit du canal est faible par rapport à la puissance de l'émetteur.

Remerciements.

Je voudrais exprimer ici ma profonde reconnaissance à MM. les Professeurs Robert Fortet et Jean Ville pour la part qu'ils ont prise à l'élaboration du présent travail, ainsi qu'à M. Charles Pisot qui m'a guidé dans la préparation de la seconde thèse. Je remercie également M. Jacques Neveu, Maître de Conférences Associé à la Faculté des Sciences de Paris, de son aide et de ses conseils.

Mon travail a été grandement facilité par l'encouragement que j'ai reçu de M. Paul Huet, Directeur de la Société *Télécommunications Radioélectriques et Téléphoniques*, pendant la durée des fonctions que j'ai exercées au Laboratoire d'Etude des Transmissions de cette Société. J'y ai reçu également un enseignement très profitable sur les questions de transmission de l'information auprès de M. Jacques Daguet, Ingénieur en Chef des Télécommunications. Qu'ils veuillent bien trouver ici l'expression de ma profonde gratitude.

BIBLIOGRAPHIE

- N. ARONSZAJN. — Theory of reproducing kernels, *Trans. Amer. Math. Soc.*, vol. 68, pp. 337-404, 1950.
- P. BETHOUX. — Tests et estimations concernant certaines fonctions aléatoires, en particulier laplaciennes, Thèse de Doctorat, Paris, 1961.
- H. CRAMER [1]. — Remarques sur le problème de prédiction pour certaines classes de processus stochastiques, *Coll. Intern. C. N. R. S.*, LXXXVII, Paris, 1958.
- [2]. — A contribution to the theory of stochastic processes, *Proc. Sec. Berkeley Symp.*, 1951, pp. 329-339.
- [3]. — On the structure of purely non-deterministic stochastic processes, *Arkiv för Mat. T.*, 4, 1961, pp. 249-266.
- J. L. DOOB. — Stochastic processes, Wiley, New York, 1953.
- U. GRENANDER & M. ROSENBLATT. — Statistical analysis of stationary time series, Wiley, New York, 1957.
- U. GRENANDER & G. SZEGÖ. — Toeplitz forms and their applications, California, 1958.

- J. M. GUELFAND & A. M. JAGLOM. — Ueber die Berechnung der Menge an Information über eine zufällige Funktion, die in einer anderen zufälligen Funktion enthalten ist, *Arbeiten zur Informationstheorie*, tome II, V. E. B., Berlin, 1958.
- T. HIDA. — Canonical representation of Gaussian processes and their application, *Mem. Coll. Science University Kyoto*, Series A, vol. XXXIII, 1, pp. 109-159, 1960.
- K. KARHUNEN [1]. — Ueber lineare Methoden in der Wahrscheinlichkeitsrechnung, *Annales Acad. Scient. Fennicae*, Serie A, 37, 1947.
- [2]. — Ueber Struktur stationärer zufälligen Funktionen, *Arkiv for Mat. T.*, tome 1, 13, 1949.
- A. KOLMOGOROV. — Suites stationnaires dans l'espace de Hilbert, *Bull. Univ. Moscou*, vol. II, 6, 1941.
- H. KOREZLIOGLU [1]. — Sur la prévision et le filtrage linéaire des processus discrets du second ordre, *Comptes Rendus A. Sc.*, tome 248, pp. 356-357, 1959.
- [2]. — Extension du théorème de Pinsker, *Comptes Rendus, A. Sc.*, tome 248, pp. 523-525, 1959.
- [3]. — Quantité d'information mutuelle par unité de temps entre deux processus aléatoires vectoriels gaussiens, stationnaires et stationnairement corrélés, *Comptes Rendus, A. Sc.*, tome 250, pp. 1436-1438, 1960.
- P. LÉVY [1]. — Wiener's random function and other Laplacian random functions, *Proc. Second Berkeley Symp.*, pp. 171-187, 1951.
- [2]. — A special problem of Brownian motion and a general theory of Gaussian random functions, *Proc. Third Berkeley Symp.*, pp. 133-177, 1955.
- [3]. — Sur une classe de courbes de l'espace de Hilbert et sur une équation intégrale non linéaire, *Annales Scient. E. N. S.*, tome 73, pp. 121-156, 1956.
- M. LOÈVE. — Probability theory. Van Nostrand, New York, 1958.
- M. S. PINSKER [1]. — Quantité d'information contenue dans un processus gaussien stationnaire par rapport à un autre processus stationnaire qui lui est stationnairement corrélé, *Doklady Akad. Nauk.*, 99, pp. 213-216, 1954.
- [2]. — Extrapolation des processus aléatoires vectoriels et quantité d'information contenue dans un processus vectoriel, stationnaire et gaussien par rapport à un autre processus qui lui est stationnairement corrélé, *Doklady Akad. Nauk.*, 121, pp. 49-52, 1958.
- A. PEREZ. — Sur la théorie de l'information dans le cas d'un alphabet abstrait, *Trans. first Prague Conf.*, pp. 209-243, 1956.
- C. SHANNON & G. WEAVER. — The mathematical theory of communication, Illinois, Urbana, 1949.
- G. VALIRON. — Cours d'analyse mathématique. — Théorie des fonctions, tome 1, Masson, Paris, 1942.
- N. VIENER & P. MASANI. — The prediction theory of multivariate stochastic processes, *Acta Math.* 98, 1-2, pp. 111-150, 1959.

DEUXIÈME THÈSE

Propositions données par la Faculté

Suites uniformément réparties
modulo un

Vu et approuvé :

Paris, le 11 janvier 1963.

Le Doyen de la Faculté des Sciences,

Marc ZAMANSKY.

Vu et permis d'imprimer :

Le Recteur de l'Académie de Paris,

Jean ROCHE.

LES IMPRESSIONS TECHNIQUES
56, Boulevard de la Villette, 56
PARIS (XIX^e)
